

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: September 17, 2004, 11:02:14 : Search time 121 Seconds
(without alignments)
200.819 Million cell updates/sec

Title: US-09-625-137-21
Perfect score: 506
Sequence: 1 KTCSPKQFACRQITCISKG.....LNGVQDCMDGSDGPHCRE 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	506	100.0	86	5	Aau81033 Human alp
2	506	100.0	790	4	Aam78716 Human pro
3	506	100.0	4544	2	Aar47861 Alpha 2-M
4	506	100.0	4544	2	Aar60517 Human alp
5	506	100.0	4544	4	Aam79091 Human pro
6	506	100.0	4544	5	Aau81019 Human alp
7	506	100.0	4544	6	Abp56839 Human LRP
8	506	100.0	4544	6	Abu89744 Protein d
9	506	100.0	4544	7	Add14025 Human src
10	506	100.0	4563	4	Abb11353 Human LBL
11	486	96.0	4529	5	Aau81016 Mouse alp
12	486	96.0	4545	5	Aau74797 Mouse alp
13	281	55.5	4636	4	Aae11937 Human Cg1
14	280	55.3	639	4	Aae11928 Human Cg1
15	280	55.3	4599	6	Abp56837 Human LRP
16	275	54.3	165	4	Abg18412 Novel hum
17	258	51.0	44	5	Aau81032 Human alp
18	253	50.0	43	5	Aau81034 Human alp
19	227	44.9	38	6	Abu61280 Human A d
20	224	44.3	37	6	Abu61281 Human A d
21	211.5	41.8	81	5	Aau81064 Human alp
22	211.5	41.8	123	5	Aau81082 Human alp
23	211.5	41.8	170	5	Aau81059 Human alp
24	211.5	41.8	209	5	Aau81056 Human alp
25	211.5	41.8	248	5	Aau81052 Human alp

26	211.5	41.8	289	5	Aau81047 Human alp
27	205.5	40.6	4601	4	Abb59371 Drosophil
28	202	39.9	727	2	Aar05533 Fragment
29	202	39.9	761	4	Abb60973 Drosophil
30	202	39.9	792	4	Abb61029 Drosophil
31	197.5	39.0	1963	4	Abb58053 Drosophil
32	194	38.3	819	6	Ada54122 Human pro
33	193	38.1	904	7	Add93401 Human lip
34	193	38.1	963	6	Abp56838 Human ado
35	193	38.1	1012	5	Aau91285 Human NOV
36	191.5	37.8	159	2	Aar05532 Fragment
37	191.5	37.8	695	5	Aau91286 Human NOV
38	191.5	37.8	699	6	Abu56579 Lung canc
39	191.5	37.8	729	5	Aau78665 Human NOV
40	191.5	37.8	729	5	Aau91282 Human NOV
41	191.5	37.8	762	5	Aau91283 Human NOV
42	191.5	37.8	762	5	Aau78666 Human NOV
43	191.5	37.8	775	7	Add93398 Human lip
44	191.5	37.8	834	7	Add93402 Human lip
45	191.5	37.8	847	5	Aau91289 Human NOV

ALIGNMENTS

RESULT 1

AAU81033

ID AAU81033 standard; protein; 86 AA.

AC AAU81033;

DT 09-APR-2002 (first entry)

DE Human alpha2 macroglobulin (alpha2M) receptor #2 peptide #2.

KW Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;
KW proliferative disorder; cancer; infectious disease; bacterial infection;
KW intracellular parasite; hypercholesterolemia; protozoan infection;
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.
XX Homo sapiens.

OS Homo sapiens.

PN WO200192474-A1.

XX 06-DEC-2001.

PD 04-JUN-2001; 2001WO-US018041.

PF 02-JUN-2000; 2000US-0209095P.

PR 25-JUL-2000; 2000US-00625137.

PR 22-SEP-2000; 2000US-00668724.

PR 28-DEC-2000; 2000US-00750972.

XX (UYCO-) UNIV CONNECTICUT HEALTH CENT.

PA Srivastava PK;

PI WPI; 2002-122061/16.

XX Screening assays for identifying compounds useful for treating immune

PT disorders, comprises identification of compounds that modulate alpha 2-

PS macroglobulin receptor-heat shock protein interaction.

XX Disclosure; Page 22; 236pp; English.

CC The invention relates to screening assays comprising identification of
CC compounds that modulate alpha2 macroglobulin (alpha2M) receptor (which
CC also functions as a heat shock protein (HSP) receptor)-HSP interaction. A
CC compound that modulates the activity of an alpha2M receptor-HSP complex
CC can be identified by contacting the compound with HSP and alpha2M

CC receptor and measuring the level of alpha2M activity or expression. If
 CC the level differs from that perceived in the absence of the test
 CC compound, a compound that modulates an alpha2M receptor-HSP-mediated
 CC process is identified. The identified compounds are useful for treating
 CC autoimmune disorders (such as multiple sclerosis or rheumatoid
 CC arthritis), diseases or disorders involving disruption of antigen
 CC presentation, endocytosis, cytokine clearance or inflammation,
 CC proliferative disorders (such as cancers including sarcomas and
 CC carcinomas), infectious diseases (such as those caused by viruses,
 CC bacteria, protozoans and intracellular parasites), hypercholesterolaemia,
 CC Alzheimer's disease, diabetes and osteoporosis. Sequences AAU81016-
 CC AAU81073 represent human and mouse alpha2M receptors and peptide
 CC fragments of the invention
 XX
 SQ Sequence 86 AA;

Query Match 100.0%; Score 506; DB 5; Length 86;
 Best Local Similarity 100.0%; Pred. No. 2.4e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60
 DB 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60
 QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
 DB 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86

RESULT 2
 AAM78716
 ID AAM78716 standard; protein; 790 AA.
 AC AAM78716;
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1378.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 OS
 XX WO200157190-A2.
 PN 09-AUG-2001.
 PD
 XX 05-FEB-2001; 2001WO-US004098.
 PF 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK51849.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX

PS Claim 20; Page 3637-3639; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 790 AA;

Query Match 100.0%; Score 506; DB 4; Length 790;
 Best Local Similarity 100.0%; Pred. No. 2.3e-42;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60
 DB 25 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 84
 QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
 DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 3
 AAR47861
 ID AAR47861 standard; protein; 4544 AA.
 AC AAR47861;
 DT 25-MAR-2003 (revised)
 DT 20-JUL-1994 (first entry)
 XX Alpha 2-Macroglobulin/LDL-receptor related protein.
 DE alpha-2 macroglobulin; Low Density Lipoprotein; LDL receptor family;
 KW LDL receptor related protein; LRP; small rhinovirus receptor; deriv;
 KW minor Rhinovirus; alpha2MR/LRP.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 211..260
 FT Misc-difference /note= "50 residues not shown in SEQ.ID.No.4"
 FT Misc-difference /note= "Residue not shown in SEQ.ID.No.4"
 FT Misc-difference /note= "Residue not shown in SEQ.ID.No.4"
 FT Misc-difference 3050
 XX WO9401553-A1.
 PN 20-JAN-1994.
 PD 05-JUL-1993; 93WO-EF001728.
 PF 08-JUL-1992; 92DE-04222385.
 PR 22-AUG-1992; 92DE-04227892.
 PR 19-FEB-1993; 93DE-04305063.
 XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA Hofer F, Kowalski H, Gruenberger M, Machat H, Huettlinger M;
 PI Kerjaeschki D, Kuechler E, Blasas D;
 XX WPI; 1994-035060/04.
 DR
 XX

PT New peptide: derivs. of receptor for rhinovirus - of the small receptor
PT gp., and derived DNA, transformed cells and antibodies, used e.g. to
PT treat or prevent rhinovirus infection.

XX Claim 5; Fig 2; 76pp; German.

XX
CC Functional derivatives of members of the Minor Rhinovirus Receptor group
CC are claimed. The alpha-2 Macroglobulin/LDL-receptor related protein of
CC sequence AAR47861 (Hetz et al. EMBO J. 7:4119-4127 (1988)) is a preferred
CC parent receptor. The derivs, which are preferably soluble, extracellular
CC forms of the native receptors, are useful for treating and preventing
CC viral (esp. rhinoviral) infections. N.B. the SEQ.ID. listing includes a
CC sequence (no.4) which differs from the alpha2-MR/LRP sequence as
CC indicated in the Features Table. (Updated on 25-MAR-2003 to correct PN
CC field.)

XX SQ Sequence 4544 AA;

Query Match 100.0%; Score 506; DB 2; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 4

AAR60517
ID AAR60517 standard; protein; 4544 AA.

XX AAR60517;

XX 25-MAR-2003 (revised)

DT 22-MAR-1995 (first entry)

XX Human alpha-2-MR.

XX Serine protease; Factor-Xa; recognition site; fusion protein cleavage;
KW protein folding; alpha-2-MR;
KW alpha-2-macroglobulin receptor/low density lipoprotein receptor.

XX Homo sapiens.

XX WO9418227-A2.

XX 18-AUG-1994.

XX 04-FEB-1994; 94WO-DK000054.

XX 04-FEB-1993; 93DK-00000130.

XX 05-FEB-1993; 93DK-00000139.

XX 03-DEC-1993; 93WO-GB002492.

XX (DENZ-) DENZYME APS.

XX Thogersen HC, Holtet TL, Etzerodt M;

XX WPI; 1994-279681/34.

XX Refolding of polypeptide molecules - using a cyclic process involving
XX denaturing and renaturing conditions to produce a correctly folded prod.

XX Disclosure; Page 131-146; 202pp; English.

XX Various domains and domain clusters of human alpha-2-MR protein have been
CC PCR amplified using the primers given in AAR71252-65. (Updated on 25-MAR-
CC 2003 to correct PN field.)

SQ Sequence 4544 AA;

Query Match 100.0%; Score 506; DB 2; Length 4544;

Best Local Similarity 100.0%; Pred. No. 1.4e-41;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 5

AAM79091

ID AAM79091 standard; protein; 4544 AA.

XX AAM79091;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1753.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00820325.

XX 01-SEP-2000; 2000US-00854936.

XX 15-SEP-2000; 2000US-00663561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

DR WPI; 2001-476283/51.

DR N-PSDB; AAK52224.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.

XX Claim 20; Page 4086-4095; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

CC sequence listing were missing at the time of publication

XX
SQ Sequence 4544 AA;
Query Match 100.0%; Score 506; DB 4; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDOITCISKWRCDEGDCPDGDEAPEICPQSKAQRCPQNEHNCIGTEL 60
DB 25 KTCSPKQFACRDOITCISKWRCDEGDCPDGDEAPEICPQSKAQRCPQNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 6

AAU81019
ID AAU81019 standard; protein; 4544 AA.

XX
AC AAU81019;

DT 09-APR-2002 (first entry)

DE Human alpha2 macroglobulin (alpha2M) receptor #2.

XX Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;
KW proliferative disorder; cancer; infectious disease; bacterial infection;
KW intracellular parasite; hypercholesterolaemia; protozoan infection;
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.

XX Homo sapiens.

XX W0200192474-A1.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018041.

XX 02-JUN-2000; 2000US-0209095P.

XX 25-JUL-2000; 2000US-00625137.

XX 22-SEP-2000; 2000US-00668724.

XX 28-DEC-2000; 2000US-00750972.

XX (UYCO-) UNIV CONNECTICUT HEALTH CENT.

XX Srivastava PK;

XX WPI; 2002-122061/16.

XX N-PSDB; ABK24096.

XX Screening assays for identifying compounds useful for treating immune

XX disorders, comprises identification of compounds that modulate alpha 2-

XX macroglobulin receptor-heat shock protein interaction.

XX Claim 8; Fig 14B; 236pp; English.

XX The invention relates to screening assays comprising identification of
CC compounds that modulate alpha2 macroglobulin (alpha2M) receptor (which
CC also functions as a heat shock protein (HSP) receptor)-HSP interaction. A
CC compound that modulates the activity of an alpha2M receptor-HSP complex
CC can be identified by contacting the compound with HSP and alpha2M
CC receptor and measuring the level of alpha2M activity or expression. If
CC the level differs from that perceived in the absence of the test
CC compound, a compound that modulates an alpha2M receptor-HSP-mediated
CC process is identified. The identified compounds are useful for treating
CC autoimmune disorders (such as multiple sclerosis or rheumatoid
CC arthritis), diseases or disorders involving disruption of antigen
CC presentation, endocytosis, cytokine clearance or inflammation,

CC proliferative disorders (such as cancers including sarcomas and
CC carcinomas), infectious diseases (such as those caused by viruses,
CC bacteria, protozoans and intracellular parasites), hypercholesterolaemia,
CC Alzheimer's disease, diabetes and osteoporosis. Sequences AAU81016-
CC AAU81073 represent human and mouse alpha2M receptors and peptide
CC fragments of the invention

XX Sequence 4544 AA;

XX Query Match 100.0%; Score 506; DB 5; Length 4544;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-41;

XX Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDOITCISKWRCDEGDCPDGDEAPEICPQSKAQRCPQNEHNCIGTEL 60
DB 25 KTCSPKQFACRDOITCISKWRCDEGDCPDGDEAPEICPQSKAQRCPQNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86

DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 7

ABP56839

ID ABP56839 standard; protein; 4544 AA.

XX
AC ABP56839;

DT 03-APR-2003 (first entry)

DE Human LRP protein SEQ ID NO:6.

XX Human; sex hormone binding globin; SHBG; MEGF-7; megalin; LRP1B;
KW low density lipoprotein receptor related protein; LDL; VLDL receptor;
KW very low density lipoprotein receptor; apolipoprotein E receptor 2;
KW cubulin; steroid hormone; steroid hormone binding protein; cytostatic;
KW contraceptive; androgen; oestrogen; progestogen; corticoid; testosterone;
KW dihydrotestosterone; oestradiol; prostate cancer; breast cancer.

XX Homo sapiens.

XX W0200299438-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-DK000379.

XX 07-JUN-2001; 2001DK-00000887.

XX 08-JUN-2001; 2001US-0296489P.

XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX (UYAA-) UNIV AARHUS.

XX Willnow T, Nykjaer A;

XX WPI; 2003-140657/13.

XX Screening for a compound that alters uptake of steroid hormone (SH) into
PT cells presenting SH binding protein receptor, by identifying compound
PT that alters uptake by the receptor of SH bound to SH binding protein.

XX Claim 21; Page 151-171; 188pp; English.

XX The present invention describes a method (M1) of screening for a compound
CC which alters uptake of steroid hormones (SH) into cells presenting a SH
CC binding protein (SBP) receptor (I). The method involves providing an
CC assay for measuring binding to or uptake by (I) of SH bound to or
CC complexed with SBP, adding the compound to be tested to the assay, and
CC determining the amount of SH bound to or complexed with SBP which is
CC bound to or taken up by (I), comparing the amount determined with an
CC amount measured in the absence of the compound to be tested, where a
CC difference in the two amounts identifies a compound which alters the
CC binding or uptake of SH bound to or complexed with SBP. Also described is

CC a compound (II) which alters the uptake of SH into cells presenting (I) as identified by (M1). (I) has cytostatic and contraceptive activities.
CC (M1) is useful for screening for a compound which alters uptake of SH.
CC such as androgens, oestrogens, progestogens and corticoids, preferably
CC androgens and oestrogens (e.g. testosterone, dihydrotestosterone, and
CC oestradiol), into cells presenting (I). (II) is useful for treating a
CC clinical condition such as prostate or breast cancer in an individual.
CC (II) is also useful for preventing pregnancy in a male. (II) is useful
CC for preparing a pharmaceutical composition for treating a clinical
CC condition associated with a steroid hormone, e.g. prostate cancer or
CC breast cancer. The present sequence represents the specifically claimed
CC human low density lipoprotein receptor related protein (LRP) protein,
CC which is given in the exemplification of the present invention
XX
SQ Sequence 4544 AA;

Query Match 100.0%; Score 506; DB 6; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 8
ID ABU89744 standard; protein; 4544 AA.
AC ABU89744;
XX
XX 10-JUL-2003 (first entry)
XX
XX Protein differentially expressed in cardiovascular disease #38.
XX
XX Cardiovascular disease; arteriosclerosis; ischaemia; angina pectoris;
KW myocardial infarction; cardiast; arteriosclerotic; antianginal;
KW gene therapy; differential gene expression.
XX
XX Homo sapiens.
XX
XX WO2003031650-A2.
PN
XX 17-APR-2003.
PD
XX 02-OCT-2002; 2002WO-EP011034.
PF
XX 08-OCT-2001; 2001GB-00024145.
PR
XX (FARB) BAYER AG.
PA
XX Munnes M, Gehrman M, Wick M, Schmitz G;
PI
XX WPI; 2003-403108/38.
DR
XX N-PSDB; ACA89917.
DR
XX Predicting, diagnosing or prognosing a cardiovascular disease, e.g.
PT angina, ischaemia, myocardial infarction or arteriosclerosis by detection
PT of a polynucleotide in a biological sample comprises detecting a
PT hybridization complex.
XX
XX Claim 3; Page 353-373; 454pp; English.
PS
XX The invention describes a method of predicting, diagnosing or prognosing
CC a cardiovascular disease by detection of a polynucleotide in a biological
CC sample comprises hybridising at least one of the polynucleotide to a
CC nucleic acid material of a biological sample, thus forming a
CC hybridisation complex, and detecting the hybridisation complex. The
CC polynucleotides, polypeptides, antisense molecule, antibody and reagent

CC are useful for preparing compositions for preventing, predicting or
CC diagnosing, or a medicament for treating a cardiovascular disease, e.g.
CC arteriosclerosis, ischaemia, angina pectoris, or myocardial infarction.
CC This sequence represents a protein identified in the invention a being
CC differentially expressed in individuals with cardiovascular disease
XX
SQ Sequence 4544 AA;

Query Match 100.0%; Score 506; DB 6; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 9
ID ADD14025 standard; protein; 4544 AA.
XX
XX ADD14025;
AC
XX 01-JAN-2004 (first entry)
DT
XX Human src biomarker polypeptide SEQ ID NO:214.
DE
XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
XX Homo sapiens.
OS
XX WO2003062395-A2.
PN
XX 31-JUL-2003.
PD
XX 17-JAN-2003; 2003WO-US001981.
PF
XX 18-JAN-2002; 2002US-0350061P.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PA
XX Huang F, Fairchild CR, Lee FY, Shaw P;
PI
XX WPI; 2003-636735/60.
DR
XX N-PSDB; ADD14619.
DR
XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
XX Claim 10; SEQ ID NO 214; 139pp; English.
PS
XX The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of

CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 4544 AA;
Query Match 100.0%; Score 506; DB 7; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDDEAPEICQSKAQCOPNEHNCGLGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDDEAPEICQSKAQCOPNEHNCGLGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGEGPHCRE 110

RESULT 10
ID ABB11353
XX AC ABB11353 standard; peptide; 4563 AA.
XX AC ABB11353;
XX DT 11-JAN-2002 (first entry)
XX DE Human LDL receptor precursor homologue, SEQ ID NO:1723.
XX KW Human; cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;
KW hematopoiesis regulation; tissue growth; immunomodulator; actin;
KW limbin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; hematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.

XX OS Homo sapiens.
XX PN WO200157188-A2.
XX XX 09-AUG-2001.
XX XX 05-FEB-2001; 2001WO-US003800.
XX XX 03-FEB-2000; 2000US-00496914.
XX XX 27-APR-2000; 2000US-00560875.
XX XX (HYSE-) HYSEQ INC.
XX XX Tang YT, Liu C, Drmanac RT;
XX PI WPI; 2001-457740/49.
XX DR N-PSDB; ABA08597.
XX PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.

XX PS Claim 20; Page 169-172; 1963pp; English.
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABB08225-ABB09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention

XX SQ Sequence 4563 AA;
Query Match 100.0%; Score 506; DB 4; Length 4563;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDDEAPEICQSKAQCOPNEHNCGLGTEL 60
DB 44 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDDEAPEICQSKAQCOPNEHNCGLGTEL 103
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 104 CVPMSRLCNGVQDCMDGSDGEGPHCRE 129

RESULT 11
ID AAU81016 standard; protein; 4529 AA.
XX AC AAU81016;
XX XX 09-APR-2002 (first entry)
XX DE Mouse alpha2 macroglobulin (alpha2M) receptor.
XX KW Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;
KW proliferative disorder; cancer; infectious disease; bacterial infection;
KW intracellular parasite; hypercholesterolemia; protozoan infection;
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.

18-DEC-2001 (first entry)
Human CGI168 (or C595) receptor protein #2.
Human, apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
cardiovascular disease; lipid metabolism; myocardial infarction;
cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
coronary artery thrombosis; cerebral artery thrombosis; stroke;
intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
neuroprotectant; cerebroprotective.
Homo sapiens.
WO200179446-A2.
25-OCT-2001.
16-APR-2001; 2001WO-US012529.
14-APR-2000; 2000US-0197137P.
20-JUN-2000; 2000US-00598042.
03-AUG-2000; 2000US-00631451.
22-SEP-2000; 2000US-00667298.
17-NOV-2000; 2000US-00714936.
(HYSE-) HYSEQ INC.
Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
Wang D;
WPI; 2001-611724/70.
N-PSDB; AAD11937.
Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases.
Claim 10; Fig 5; 266pp; English.
The invention relates to polynucleotides encoding proteins CGI122, CGI179,
CG95, CGI121, CGI162, CG27, CGI153 and CGI168 which are related to proteins
involved in lipid metabolism and cardiovascular disease such as human
apolipoproteins, lipases and lipoprotein receptor proteins. These DNA and
protein sequences are useful for treating or preventing disorders
associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
expression and for treating lipid metabolism, cardiovascular diseases and
thrombosis. Antibodies against these proteins are useful for determining
the presence of or predisposition to a disease associated with altered
levels of these sequences. ALR polypeptides are also useful for
identifying agents (agonists and antagonists) that bind to them and cells
expressing ALR proteins are useful for identifying a therapeutic agent
for use in treatment of a pathology related to aberrant expression or
physiological interactions of this polypeptide. Vectors comprising these
DNA and protein sequences are also useful for producing ALR proteins.
The nucleic acids and polypeptides of the invention are also useful for
the treatment of occlusive cardiovascular diseases, myocardial
infarction, cerebral ischaemia, angina, arterial thrombosis, coronary
artery thrombosis and cerebral artery thrombosis or intracardiac
thrombosis and stroke. The nucleotides of the invention are used in gene
therapy. The present sequence is human CGI168 (or C595) receptor protein
Sequence 4636 AA;
Query Match 55.5%; Score 281; DB 4; Length 4636;
Best Local Similarity 53.8%; Pred. No. 5.5e-19;
Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;
QY 3 CSPKQFACRQITICISKGRNCDGSDPDSAPICQSAQRQCPNHNCLGTCLCV 62
DB 69 CDFGEFLCHDVTVCVSRSMLCDGDPDSDSLTCTPCEVEIKCPNLHACLGNKCV 128
QY 63 PMSRLCNGVQDCMDGSDGPHCRE 86

Db 129 HSLQCNGLVDCPDGDEGVHQCQ 152
RESULT 14
AAE11928
ID AAE11928 standard; protein; 639 AA.
XX
AC AAE11928;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human CGI168 (or C595) receptor protein #1.
XX
KW Human, apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
cardiovascular disease; lipid metabolism; myocardial infarction;
cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
coronary artery thrombosis; cerebral artery thrombosis; stroke;
intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
neuroprotectant; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200179446-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US012529.
XX
PR 14-APR-2000; 2000US-0197137P.
PR 20-JUN-2000; 2000US-00598042.
PR 03-AUG-2000; 2000US-00631451.
PR 22-SEP-2000; 2000US-00667298.
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
DR WPI; 2001-611724/70.
DR N-PSDB; AAD19223.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
receptor polypeptides useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases.
XX
PS Claim 10; Fig 5; 266pp; English.
XX
CC The invention relates to polynucleotides encoding proteins CGI122, CGI179,
CG95, CGI121, CGI162, CG27, CGI153 and CGI168 which are related to proteins
involved in lipid metabolism and cardiovascular disease such as human
apolipoproteins, lipases and lipoprotein receptor proteins. These DNA and
protein sequences are useful for treating or preventing disorders
associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
expression and for treating lipid metabolism, cardiovascular diseases and
thrombosis. Antibodies against these proteins are useful for determining
the presence of or predisposition to a disease associated with altered
levels of these sequences. ALR polypeptides are also useful for
identifying agents (agonists and antagonists) that bind to them and cells
expressing ALR proteins are useful for identifying a therapeutic agent
for use in treatment of a pathology related to aberrant expression or
physiological interactions of this polypeptide. Vectors comprising these
DNA and protein sequences are also useful for producing ALR proteins.
The nucleic acids and polypeptides of the invention are also useful for
the treatment of occlusive cardiovascular diseases, myocardial
infarction, cerebral ischaemia, angina, arterial thrombosis, coronary
artery thrombosis and cerebral artery thrombosis or intracardiac
thrombosis and stroke. The nucleotides of the invention are used in gene
therapy. The present sequence is human CGI168 (or C595) receptor protein
Sequence 639 AA;
QY

Query Match 55.3%; Score 280; DB 4; Length 639;
 Best Local Similarity 53.6%; Pred. No. 9.2e-20;
 Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

CC androgens and oestrogens (e.g. testosterone, dihydrotestosterone, and
 CC oestradiol), into cells presenting (I). (II) is useful for treating a
 CC clinical condition such as prostate or breast cancer in an individual.
 CC (II) is also useful for preventing pregnancy in a male. (II) is useful
 CC for preparing a pharmaceutical composition for treating a clinical
 CC condition associated with a steroid hormone, e.g. prostate cancer or
 CC breast cancer. The present sequence represents the specifically claimed
 CC human low density lipoprotein receptor related protein 1B (LRP1B)
 CC protein, which is given in the exemplification of the present invention
 XX

QY 3 CSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPSKAQRCPNEHNCIGTGLCV 62
 Db 69 CDGFEFLCHDHVTCVQSQWLCDGDPDSDSDSLDTCPEEVEIKPLNHIACLGTKVCV 128
 QY 63 PMSRLNGVQDCMDGSDGPHCRE 86
 Db 129 HLSQLCNGVLDCPDGYDEGVHCQE 152

Query Match 55.3%; Score 280; DB 6; Length 4599;
 Best Local Similarity 53.6%; Pred. No. 6.9e-19;
 Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPSKAQRCPNEHNCIGTGLCV 62
 Db 32 CDGFEFLCHDHVTCVQSQWLCDGDPDSDSDSLDTCPEEVEIKPLNHIACLGTKVCV 91
 QY 63 PMSRLNGVQDCMDGSDGPHCRE 86
 Db 92 HLSQLCNGVLDCPDGYDEGVHCQE 115

Search completed: September 17, 2004, 11:15:23
 Job time : 124 secs

RESULT 15
 ABP56837
 ID ABP56837 standard; protein; 4599 AA.
 XX
 AC ABP56837;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE Human LRP1B protein SEQ ID NO:4.
 XX
 KW Human; sex hormone binding globin; SHBG; MEGF-7; megalin; LRP1B;
 KW low density lipoprotein receptor related protein; LDL; VLDL receptor;
 KW very low density lipoprotein receptor; apolipoprotein E receptor 2;
 KW cubulin; steroid hormone; steroid hormone binding protein; cytosolic;
 KW contraceptive; androgen; oestrogen; progesterone; corticoid; testosterone;
 KW dihydrotestosterone; oestradiol; prostate cancer; breast cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 4343
 FT /note= "any amino acid"
 XX
 PN WO200299438-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-DK000379.
 XX
 PR 07-JUN-2001; 2001DK-00000887.
 PR 08-JUN-2001; 2001US-0296489P.
 XX
 PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
 PA (UYAA-) UNIV AARHUS.
 XX
 PI Willnow T, Nykjaer A;
 XX
 DR WPI; 2003-140657/13.
 XX
 PT Screening for a compound that alters uptake of steroid hormone (SH) into
 PT cells presenting SH binding protein receptor, by identifying compound
 PT that alters uptake by the receptor of SH bound to SH binding protein.
 XX
 PS Claim 21; Page 126-146; 188pp; English.
 XX
 CC The present invention describes a method (M1) of screening for a compound
 CC which alters uptake of steroid hormones (SH) into cells presenting a SH
 CC binding protein (SBP) receptor (I). The method involves providing an
 CC assay for measuring binding to or uptake by (I) of SH bound to or
 CC complexed with SBP, adding the compound to be tested to the assay, and
 CC determining the amount of SH bound to or complexed with SBP which is
 CC bound to or taken up by (I), comparing the amount determined with an
 CC amount measured in the absence of the compound to be tested, where a
 CC difference in the two amounts identifies a compound which alters the
 CC binding or uptake of SH bound to or complexed with SBP. Also described is
 CC a compound (II) which alters the uptake of SH into cells presenting (I)
 CC as identified by (M1). (I) has cytostatic and contraceptive activities.
 CC (M1) is useful for screening for a compound which alters uptake of SH
 CC such as androgens, oestrogens, progesterone and corticoids, preferably

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:13:19 ; Search time 32 Seconds
(without alignments)
138.745 Million cell updates/sec

Title: US-09-625-137-21

Perfect score: 506

Sequence: 1 KTCSPKQFACRDQITCSKG.....LCNGVQDCMGDSGEPHCRC 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/PCUTS_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	4544	1	US-08-469-486-52
2	506	100.0	4544	2	US-08-469-558-52
3	202	39.9	726	6	5208144-37
4	191.5	37.8	159	6	5208144-35
5	187.5	37.1	2214	1	US-08-727-034-7
6	184.5	36.5	2213	1	US-08-727-034-3
7	183	36.2	4654	3	US-08-476-515A-84
8	183	36.2	4655	3	US-08-652-877-84
9	183	36.2	4655	3	US-08-652-877-88
10	183	36.2	4655	3	US-08-652-877-90
11	182	36.0	4655	3	US-08-652-877-86
12	180.5	35.7	846	1	US-08-149-103-3
13	180.5	35.7	846	1	US-08-149-103-4
14	180.5	35.7	846	1	US-08-451-883-3
15	180.5	35.7	846	1	US-08-451-883-4
16	180.5	35.7	873	1	US-08-393-734-2
17	180.5	35.7	873	3	US-08-894-489-2
18	176.5	34.9	884	6	5208144-8
19	176	34.8	356	1	US-08-228-162-2
20	176	34.8	860	1	US-08-092-817-4
21	176	34.8	860	4	US-08-485-128-4
22	176	34.8	860	4	US-09-804-778A-8
23	176	34.8	1074	2	US-08-470-058-2
24	176	34.8	1074	3	US-09-037-188-2
25	176	34.8	1074	3	US-09-285-310-2
26	176	34.8	1410	2	US-08-470-058-4
27	176	34.8	1410	3	US-09-037-188-4

28 176 34.8 1410 3 US-09-285-310-4 Sequence 4, Appli
29 174.5 34.5 943 3 US-08-476-515A-12 Sequence 12, Appli
30 174.5 34.5 944 3 US-08-652-877-12 Sequence 12, Appli
31 171.5 33.9 1451 4 US-09-060-299-25 Sequence 25, Appli
32 171.5 33.9 1451 4 US-09-402-923A-25 Sequence 25, Appli
33 171.5 33.9 1584 4 US-09-060-299-39 Sequence 39, Appli
34 171.5 33.9 1584 4 US-09-402-923A-39 Sequence 39, Appli
35 171.5 33.9 1591 4 US-09-060-299-4 Sequence 4, Appli
36 171.5 33.9 1591 4 US-09-060-299-43 Sequence 43, Appli
37 171.5 33.9 1591 4 US-09-402-923A-4 Sequence 4, Appli
38 171.5 33.9 1591 4 US-09-402-923A-43 Sequence 43, Appli
39 171.5 33.9 1615 4 US-09-060-299-3 Sequence 3, Appli
40 171.5 33.9 1615 4 US-09-402-923A-3 Sequence 3, Appli
41 171.5 33.9 1639 4 US-09-060-299-29 Sequence 29, Appli
42 171.5 33.9 1639 4 US-09-402-923A-29 Sequence 29, Appli
43 170.5 33.7 1586 4 US-09-060-299-44 Sequence 44, Appli
44 170.5 33.7 1586 4 US-09-402-923A-44 Sequence 44, Appli
45 170.5 33.7 1614 4 US-09-060-299-42 Sequence 42, Appli

ALIGNMENTS

RESULT 1
US-08-469-486-52
; Sequence 52, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Ezerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4544 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-486-52

Query Match 100.0%; Score 506; DB 1; Length 4544;
Best Local Similarity 100.0%; Pred. No. 8.6e-45;

```
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 2
US-08-469-658-52
; Sequence 52, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF
; PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4544 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-52

Query Match 100.0%; Score 506; DB 2; Length 4544;
Best Local Similarity 100.0%; Pred. No. 8.6e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110
```

```
RESULT 3
5208144-37
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:37:
; LENGTH: 726
5208144-37

Query Match 39.9%; Score 202; DB 6; Length 726;
Best Local Similarity 45.3%; Pred. No. 1.8e-13;
Matches 39; Conservative 11; Mismatches 32; Indels 4; Gaps 3;
QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 60
DB 397 KTCSPFTHFLC-DNGNCIYKAWICDGDNDRCRMSDE--KDCP-TQPHPCPSTQWQCGYST 452
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 453 CINLSALCDGVFCPCNGTDESPLCNQ 478

RESULT 4
5208144-35
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:35:
; LENGTH: 159
5208144-35

Query Match 37.8%; Score 191.5; DB 6; Length 159;
Best Local Similarity 41.0%; Pred. No. 4.4e-13;
Matches 34; Conservative 10; Mismatches 36; Indels 3; Gaps 1;
QY 2 TCSPKQFACRDQITCISKWRCDEGDCPDGSDAEAPICPSKAQRCQPNHNCGLGTEL 61
DB 2 TCPDPFTKCOPTNTICVPRFLCDGDNDCDGDSDENPIYC---ASHTCRSNEFQCLSPQRC 58
QY 62 VPMMSRLCNGVQDCMDGSDGPHCRE 84
DB 59 IPSYWFCDGEADCADGSDGDEPDTG 81

RESULT 5
US-08-727-034-7
; Sequence 7, Application US/08727034
; Patent No. 5655872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
```

; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,034
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 536
; APPLICATION NUMBER: JP 261440/1995
; FILING DATE: 09-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 102451/1996
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-079-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-727-034-7
;
; Query Match 37.1%; Score 187.5; DB 1; Length 2214;
; Best Local Similarity 47.6%; Pred. No. 2.1e-11;
; Matches 40; Conservative 9; Mismatches 24; Indels 11; Gaps 6;
;
; QY 2 TCSPKQFACRDQITCISKGRCDGRCDEPDCGSDPEAIPCQS-KAQRCPNHNCLGTGL 60
; Db 1198 TCASNFQCRNG-HCIPQRWACDGTDCQDGSDEPVCCKKNGFRC-PN-----GT-- 1248
;
; QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84
; Db 1249 CIPSSKHCDGLRDCSDGSDP-QHC 1271
;
; RESULT 6
; US-08-727-034-3
; Sequence 3, Application US/08727034
; Patent No. 5665872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; TITLE OF INVENTION: THE GENE CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,034
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 536
; APPLICATION NUMBER: JP 261440/1995
; FILING DATE: 09-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 102451/1996
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-079-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-727-034-3
;
; Query Match 36.5%; Score 184.5; DB 1; Length 2213;
; Best Local Similarity 46.4%; Pred. No. 4.4e-11;
; Matches 39; Conservative 9; Mismatches 25; Indels 11; Gaps 6;
;
; QY 2 TCSPKQFACRDQITCISKGRCDGRCDEPDCGSDPEAIPCQS-KAQRCPNHNCLGTGL 60
; Db 1197 TCASNFQCRNG-HCIPQRWACDGTDCQDGSDEPVCCKKNGFRC-PN-----GT-- 1247
;
; QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84
; Db 1248 CIPSTKHCDGLRDCSDGSDP-QHC 1270
;
; RESULT 7
; US-08-476-515A-84
; Sequence 84, Application US/08476515A
; Patent No. 6239270
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Tuhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjaln, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thersof and DNA Encoding Same
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Martin Savitzky
; STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.;
; STREET: 3C43,
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Compaq PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 7.0 (Patentin)
; CURRENT APPLICATION DATA:


```
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4655 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-652-877-88

Query Match 36.2%; Score 183; DB 3; Length 4655;
Best Local Similarity 42.9%; Pred. No. 1.4e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRDQITCISKGRCDGRDCPDGSDRAPEICPQSKAQCQPNHNCIGTEL 60
Db 1269 KTCPSYFHC-DNGNCIHRWLCDRDNDGDSDE--KDCP-TQPFRCPSWQWQCLGHNI 1324

QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84
Db 1325 CYNLSVVCDFGFCPNGTDESPLC 1348

RESULT 10
; Sequence 90, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjaln, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 86:
```

```
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4655 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-652-877-90

Query Match 36.2%; Score 183; DB 3; Length 4655;
Best Local Similarity 42.9%; Pred. No. 1.4e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRDQITCISKGRCDGRDCPDGSDRAPEICPQSKAQCQPNHNCIGTEL 60
Db 1269 KTCPSYFHC-DNGNCIHRWLCDRDNDGDSDE--KDCP-TQPFRCPSWQWQCLGHNI 1324

QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84
Db 1325 CYNLSVVCDFGFCPNGTDESPLC 1348

RESULT 11
; Sequence 86, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjaln, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 86:
```

SEQUENCE CHARACTERISTICS:
LENGTH: 4655 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-877-86

Query Match 36.0%; Score 182; DB 3; Length 4655;

Best Local Similarity 44.3%; Pred. No. 1.8e-10; Indels 6; Gaps 4;
Matches 35; Conservative 14; Mismatches 24;

QY 2 TCSKQFACRDQITCISKWRCDCGERDCPDGSDPEAFCPSKAQRCQPNHNCLGTCLC 61
DB 1108 SCLDTQVTC-DNHQCISKWVCDTNDGCGSDE--KNC--NSTETCQPSQFNC-PNHRC 1161

QY 62 VPMRLCNGVQDCMDGSDE 80

DB 1162 IDLSFVCDGDKDCVDGSDE 1180

RESULT 12

US-08-149-103-3

; Sequence 3, Application US/08149103

; Patent No. 5750367

; GENERAL INFORMATION:

; APPLICANT: Lawrence C. B. Chan

; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW

; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS

; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH

; TITLE OF INVENTION: RECEPTORS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LYON & LYON

; STREET: 611 West Sixth Street

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90017

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/149,103

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA: including application

; PRIOR APPLICATION DATA: described below:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 204/052

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 846 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-149-103-3

Query Match

Best Local Similarity 35.7%; Score 180.5; DB 1; Length 846;

US-08-451-883-3

Matches 35; Conservative 11; Mismatches 25; Indels 7; Gaps 4;

QY 3 CSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAFCPSKAQRCQPNHNCLGTCLC 62

DB 212 CPASEIQCGSG-ECIHKKWRCDCGDPCKDGSDEVN--CP---SRTCRPDQFECEDGS-CI 264
QY 63 PMSRLCNGVQDCMDGSDE 80
DB 265 HGSRCQNGIRDVCVDGSDE 282

RESULT 13

US-08-149-103-4

; Sequence 4, Application US/08149103

; Patent No. 5750367

; GENERAL INFORMATION:

; APPLICANT: Lawrence C. B. Chan

; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW

; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS

; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH

; TITLE OF INVENTION: RECEPTORS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LYON & LYON

; STREET: 611 West Sixth Street

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90017

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/149,103

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA: including application

; PRIOR APPLICATION DATA: described below:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 204/052

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 846 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-149-103-4

Query Match

Best Local Similarity 35.7%; Score 180.5; DB 1; Length 846;

US-08-451-883-3

Matches 35; Conservative 11; Mismatches 25; Indels 7; Gaps 4;

QY 3 CSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAFCPSKAQRCQPNHNCLGTCLC 62

DB 212 CPASEIQCGSG-ECIHKKWRCDCGDPCKDGSDEVN--CP---SRTCRPDQFECEDGS-CI 264

QY 63 PMSRLCNGVQDCMDGSDE 80

DB 265 HGSRCQNGIRDVCVDGSDE 282

RESULT 14

US-08-451-883-3

; Sequence 3, Application US/08451883

; Patent No. 5798209


```

; GENERAL INFORMATION:
; APPLICANT: Lawrence C.B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR
; TITLE OF INVENTION: USE OF SUCH RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street, Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 6.22)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,883
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/149,103
; FILING DATE: No. 5798209ember 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 212/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 846 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-451-883-3
;
; Query Match 35.7%; Score 180.5; DB 1; Length 846;
; Best Local Similarity 44.9%; Pred. No. 4e-11;
; Matches 35; Conservative 11; Mismatches 25; Indels 7; Gaps 4;
;
; QY 3 CSPKQFACRDOITCISKWRCDCGERDCPDGSDAEAPICPQSKAORCOPNEHNCIGTCLCV 62
; Db 212 CPASEIQGSG-ECIHKKWRCDCGPDCKDGSDEVN--CP---SRTCRPDQFECEDGS-CI 264
;
; QY 63 PMSRLCNGVQDCMDGSDE 80
; Db 265 HGSRCQNGIRDCVDSDE 282
;
; Search completed: September 17, 2004, 11:19:15
; Job time : 33 secs
;
;
; GENERAL INFORMATION:
; APPLICANT: Lawrence C.B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR
; TITLE OF INVENTION: USE OF SUCH RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street, Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 6.22)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,883
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/149,103
; FILING DATE: No. 5798209ember 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 212/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 846 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-451-883-3
;
; Query Match 35.7%; Score 180.5; DB 1; Length 846;
; Best Local Similarity 44.9%; Pred. No. 4e-11;
; Matches 35; Conservative 11; Mismatches 25; Indels 7; Gaps 4;
;
; QY 3 CSPKQFACRDOITCISKWRCDCGERDCPDGSDAEAPICPQSKAORCOPNEHNCIGTCLCV 62
; Db 212 CPASEIQGSG-ECIHKKWRCDCGPDCKDGSDEVN--CP---SRTCRPDQFECEDGS-CI 264
;
; QY 63 PMSRLCNGVQDCMDGSDE 80
; Db 265 HGSRCQNGIRDCVDSDE 282
;
; RESULT 15
; US-08-451-883-4
; Sequence 4, Application US/08451883
; Patent No. 5798209
; GENERAL INFORMATION:
; APPLICANT: Lawrence C.B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR
; TITLE OF INVENTION: USE OF SUCH RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street, Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
;
; one
```

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:15:30 ; Search time 48 Seconds
(without alignments)
575.364 Million cell updates/sec

Title: US-09-625-137-21
Perfect score: 506
Sequence: 1 KTCSFKQFACRQDITCISKG.....LNGVQDCMDGSDGPHCRE 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues
Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	86	12	US-09-750-972-21
2	506	100.0	2641	16	US-10-741-601-333
3	506	100.0	4183	12	US-10-087-192-672
4	506	100.0	4485	16	US-10-741-601-332
5	506	100.0	4544	16	US-10-464-368-68
6	506	100.0	4563	12	US-10-276-774-1723
7	486	96.0	3197	12	US-10-087-192-669
8	486	96.0	4545	9	US-09-873-403-2
9	486	96.0	4545	12	US-09-750-972-2
10	486	96.0	4545	16	US-10-464-368-67
11	486	96.0	4545	16	US-10-464-368-71
12	281	55.5	4636	9	US-09-835-996A-33
13	280	55.3	6339	9	US-09-835-996A-15
14	270	53.4	4599	16	US-10-464-368-69
15	270	53.4	4599	16	US-10-464-368-70

16	258	51.0	44	12	US-09-750-972-20	Sequence 20, Appl
17	253	50.0	43	12	US-09-750-972-22	Sequence 22, Appl
18	227	44.9	38	14	US-10-133-128-78	Sequence 78, Appl
19	227	44.9	38	14	US-10-289-660-78	Sequence 78, Appl
20	227	44.9	38	16	US-10-693-057-78	Sequence 79, Appl
21	224	44.3	37	14	US-10-133-128-79	Sequence 79, Appl
22	224	44.3	37	14	US-10-289-660-79	Sequence 79, Appl
23	224	44.3	37	16	US-10-693-057-79	Sequence 79, Appl
24	212	41.9	99	16	US-10-693-057-417	Sequence 417, Appl
25	211.5	41.8	81	12	US-09-750-972-52	Sequence 52, Appl
26	211.5	41.8	99	16	US-10-693-057-421	Sequence 421, Appl
27	211.5	41.8	123	12	US-09-750-972-50	Sequence 50, Appl
28	211.5	41.8	170	12	US-09-750-972-47	Sequence 47, Appl
29	211.5	41.8	209	12	US-09-750-972-44	Sequence 44, Appl
30	211.5	41.8	248	12	US-09-750-972-40	Sequence 40, Appl
31	211.5	41.8	289	12	US-09-750-972-35	Sequence 35, Appl
32	210	41.5	90	16	US-10-693-057-422	Sequence 422, App
33	204.5	40.4	4753	15	US-10-369-493-5119	Sequence 5119, Ap
34	202	39.9	4660	16	US-10-464-368-74	Sequence 74, Appl
35	201.5	39.8	90	16	US-10-693-057-425	Sequence 425, App
36	199	38.3	97	16	US-10-693-057-434	Sequence 434, App
37	194	38.3	819	15	US-10-094-749-1690	Sequence 1690, Ap
38	193	38.1	92	16	US-10-693-057-416	Sequence 416, App
39	193	38.1	963	16	US-10-464-368-86	Sequence 86, Appl
40	192.5	38.0	92	16	US-10-693-057-432	Sequence 432, App
41	191.5	37.8	2643	15	US-10-369-493-5010	Sequence 5010, Ap
42	190.5	37.6	862	14	US-10-281-478-3	Sequence 3, Appli
43	190.5	37.6	862	16	US-10-464-368-90	Sequence 90, Appl
44	190.5	37.6	862	16	US-10-464-368-91	Sequence 91, Appl
45	189	37.4	2180	15	US-10-369-493-5009	Sequence 5009, Ap

ALIGNMENTS

RESULT 1
US-09-750-972-21
; Sequence 21, Application US/09750972
; Publication No. US20040072993A1
; GENERAL INFORMATION:
; APPLICANT: Pramod K. Srivastava
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
; FILE REFERENCE: 8449-134
; CURRENT APPLICATION NUMBER: US/09/750,972
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/750,972
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/668,724
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-750-972-21

Query Match	100.0%	Score 506,	DB 12;	Length 86;
Best Local Similarity	100.0%	Pred. No. 2.1e-43;		
Matches	86;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	KTCSFKQFACRQDITCISKGWRCDGDCPDGSDGPHCRE	16	Sequence 60
Db	1	KTCSFKQFACRQDITCISKGWRCDGDCPDGSDGPHCRE	16	Sequence 60
QY	61	CVPMRLNGVQDCMDGSDGPHCRE	86	Sequence 86
Db	61	CVPMRLNGVQDCMDGSDGPHCRE	86	Sequence 86

RESULT 2
US-10-741-601-333

```
; Sequence 333, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 333
; LENGTH: 2641
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-333

Query Match      100.0%; Score 506; DB 16; Length 2641;
Best Local Similarity 100.0%; Pred. No. 6.3e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 60
DB 25 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 3
US-10-087-192-672
; Sequence 672, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 672
; LENGTH: 4183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-672

Query Match      100.0%; Score 506; DB 12; Length 4183;
Best Local Similarity 100.0%; Pred. No. 1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 60
DB 25 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 4
US-10-741-601-332
; Sequence 332, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 4485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-332

Query Match      100.0%; Score 506; DB 16; Length 4485;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 60
DB 25 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 5
US-10-464-368-68
; Sequence 68, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT FILING DATE: 2003-06-16
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 68
; LENGTH: 4544
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-68

Query Match      100.0%; Score 506; DB 16; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 60
DB 25 KTCSPKQFACRQDQITCISKWRCGDERDCPDGSDAEAPICPSKAQRCQPNHNCLGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 6
US-10-276-774-1723
; Sequence 1723, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT FILING DATE: 2002-11-18
; PRIOR FILING DATE: 2000-04-27
```

; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1723
; LENGTH: 4563
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1723

Query Match 100.0%; Score 506; DB 12; Length 4563;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60
DB 44 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 103

QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 104 CVPMSRLCNGVQDCMDGSDGEGPHCRE 129

RESULT 7
US-10-087-192-669
; Sequence 669, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669
; LENGTH: 3197
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-669

Query Match 96.0%; Score 486; DB 12; Length 3197;
Best Local Similarity 95.3%; Pred. No. 7.8e-40;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85

QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 8
US-09-873-403-2
; Sequence 2, Application US/09873403
; Patent No. US2002028207A1
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Pramod K.
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
; FILE REFERENCE: 8449-178
; CURRENT APPLICATION NUMBER: US/09/873,403
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/625,139
; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: 60/209,266
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-403-2

Query Match 96.0%; Score 486; DB 9; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85

QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 9
US-09-750-972-2
; Sequence 2, Application US/09750972
; Publication No. US20040072993A1
; GENERAL INFORMATION:
; APPLICANT: Pramod K. Srivastava
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
; FILE REFERENCE: 8449-134
; CURRENT APPLICATION NUMBER: US/09/750,972
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/750,972
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/668,724
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-750-972-2

Query Match 96.0%; Score 486; DB 12; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85

QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 10
US-10-464-368-67
; Sequence 67, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1P-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-08-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-67

Query Match          96.0%; Score 486; DB 16; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDOITCISKWRCGDERDCPDGSDPEAIPCQSKAQRCPQNEHNCILGTEL 60
DB 26 KTCSPKQFACRDOITCISKWRCGDERDCPDGSDPEAIPCQSKAQRCPQNEHNCILGTEL 85

QY 61 CVPMSRLCNGVQDCMGSDGEGPHCRE 86
DB 86 CVPMSRLCNGIQDCMGSDGEGAHCRE 111

RESULT 11
US-10-464-368-71
; Sequence 71, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40718-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-71

Query Match          96.0%; Score 486; DB 16; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDOITCISKWRCGDERDCPDGSDPEAIPCQSKAQRCPQNEHNCILGTEL 60
DB 26 KTCSPKQFACRDOITCISKWRCGDERDCPDGSDPEAIPCQSKAQRCPQNEHNCILGTEL 85

QY 61 CVPMSRLCNGVQDCMGSDGEGPHCRE 86
DB 86 CVPMSRLCNGIQDCMGSDGEGAHCRE 111

RESULT 12
US-09-835-996A-33
; Sequence 33, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunrui
```

```
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 4636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-33

Query Match          55.5%; Score 281; DB 9; Length 4636;
Best Local Similarity 53.6%; Pred. No. 4.4e-19;
Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQFACRDOITCISKWRCGDERDCPDGSDPEAIPCQSKAQRCPQNEHNCILGTEL 62
DB 69 CDPGEFLCHDVTCVSRSLWLCGDPDCPDSDSLDTCPEVEIKCPNLHIACLGINKCV 128

QY 63 PMSRLCNGVQDCMGSDGEGPHCRE 86
DB 129 HLSQLCNGVLDCPDGDEGVHQCQE 152

RESULT 13
US-09-835-996A-15
; Sequence 15, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 639
; TYPE: PRT
```

Search completed: September 17, 2004, 11:20:08
Job time : 49 secs

```

US-10-464-368-69
/ Sequence 69, Application US/10464368
/ Publication No. US20040023356A1
/ GENERAL INFORMATION:
/ APPLICANT: Krumlauf, Robb
/ APPLICANT: Ellies, Debra
/ TITLE OF INVENTION, COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
/ FILE REFERENCE: 40716-IP-017
/ CURRENT APPLICATION NUMBER: US/10/464,368
/ CURRENT FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: 60/388,970
/ PRIOR FILING DATE: 2002-06-14
/ NUMBER OF SEQ ID NOS: 140
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 69
/ LENGTH: 4599
/ TYPE: PRT
/ ORGANISM: MOUSE
US-10-464-368-69

```

```

RESULT 15
US-10-464-368-70
; Sequence 70, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumihauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1p-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 70
; LENGTH: 4599
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-70

```

Query Match 53.4%; Score 270; DB 16; Length 4599;
Best Local Similarity 51.2%; Pred. No. 5.6e-18;

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:11:04 ; Search time 39 Seconds
(without alignments)
212.115 Million cell updates/sec

Title: US-09-625-137-21

Perfect score: 506

Sequence: 1 KTCSPKQFACRDQITCISKG.....LCNGVQDCMGSDGPHCRE 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:**

1: Pirl:**

2: Pirl:**

3: Pirl:**

4: Pirl:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	4544	1 S02392	alpha-2-macroglobu
2	486	96.0	4545	1 S25111	alpha-2-macroglobu
3	459	90.7	4543	1 A53102	alpha-2-macroglobu
4	204.5	40.4	4753	1 A47437	LDL-receptor-relat
5	202	39.9	4660	2 T42737	gp330 protein prec
6	200.5	39.6	909	1 QRXL22	LDL receptor 2 pre
7	197.5	39.0	990	2 H88733	protein F32E10.3 l
8	197.5	39.0	1984	2 T13171	probable vitellogge
9	196	38.7	972	2 A30363	glycoprotein Gp330
10	192.5	38.0	909	1 QRXL21	LDL receptor 1 pre
11	191.5	37.8	879	1 QRRTLD	LDL receptor prote
12	191.5	37.8	2643	2 T29149	hypothetical prote
13	190.5	37.6	862	1 QRMSLD	LDL receptor prote
14	189	37.4	854	1 QRHYLD	LDL receptor prote
15	189	37.4	2180	2 T29764	hypothetical prote
16	187.5	37.1	1650	2 S53457	dominant autoantig
17	187	37.0	996	2 JEO237	apolipoprotein E r
18	185.5	36.7	869	1 JCA4858	VLDL receptor prec
19	185	36.6	1160	2 F88369	protein unc-52 (im
20	185	36.6	2295	2 C88369	protein unc-52 (im
21	185	36.6	3375	2 T19821	hypothetical prote
22	184.5	36.5	873	1 T48952	VLDL receptor prec
23	182.5	36.1	837	1 A29512	LDL receptor prote
24	182.5	36.1	863	1 S51789	VLDL receptor prec
25	182.5	36.1	873	1 QRRTVD	VLDL receptor prec
26	180.5	35.7	873	1 A49729	LDL receptor prote
27	177.5	35.1	2215	2 T00348	LR11 protein - mou
28	176	34.8	860	1 QRHULD	LDL receptor prote
29	176	34.8	1357	2 T16860	hypothetical prote

RESULT 1

S02392

alpha-2-macroglobulin receptor precursor - human

N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein recept

C:Species: Homo sapiens (man)

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 22-Jun-1999

C:Accession: S02392; S30027; I37998; A39210; S12538

R:Herz, J.; Hamann, U.; Rogne, S.; Myklebost, O.; Gausepohl, H.; Stanley, K.K.

EMBO J. 7, 4119-4127, 1988

A>Title: Surface location and high affinity for calcium of a 500-kd liver membrane prot

A:Reference number: S02392; MUID:89210795; PMID:3266596

A:Accession: S02392

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-4544 <HER>

A:Cross-references: ENBL:X13916; NID:G34338; PIDN:CAA32112.1; PID:G34339

R:Kristensen, T.

submitted to the EMBL Data Library, October 1990

A:Reference number: S30027

A:Accession: S30027

A:Molecule type: mRNA

A:Residues: 3275-3864 <KRI>

A:Cross-references: ENBL:X55077

R:Herz, J.; Kowal, R.C.; Goldstein, J.L.; Brown, M.S.

EMBO J. 9, 1769-1776, 1990

A>Title: Proteolytic processing of the 600 kd low density lipoprotein receptor-related

A:Reference number: S12538; MUID:90269210; PMID:2112085

A:Contents: annotation; site of proteolytic cleavage

R:Kutt, H.; Herz, J.; Stanley, K.K.

Biochim. Biophys. Acta 1009, 229-236, 1989

A>Title: Structure of the low-density lipoprotein receptor-related protein (LRP) promot

A:Reference number: I37998; MUID:90089395; PMID:2597675

A:Accession: I37998

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-11 <RES>

A:Cross-references: ENBL:X15424; NID:G34408; PIDN:CAA33464.1; PID:G34409

R:Strickland, D.K.; Ashcom, J.D.; Williams, S.; Burgess, W.H.; Migliorini, M.; Argraves

J. Biol. Chem. 265, 17401-17404, 1990

A>Title: Sequence identity between the alpha2-macroglobulin receptor and low density li

A:Reference number: A39210; MUID:91009181; PMID:1698775

A:Accession: A39210

A>Status: preliminary

A:Molecule type: protein

A:Residues: 150-166;234-238, 'X', 240-245, 'X', 247-252, 'G', 686-695;902-916;1096-1109, 'S', 'I'

C:Genetics:

A:Gene: GDB:LRP1; APR; LRP; A2MR

A:Cross-references: GDB:I19694; OMIM:107770

A:Map position: 12q13.1-12q13.3

C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associat

ciated protein (see PIR:A39875).

C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding;

low density lipopr
parlecan precursor
low density lipopr
probable vitellogge
head-activator bin
low density lipopr
heparan sulfate pr
protein B0244.8 [i
model protein prec
hypothetical prote
hypothetical prote
hypothetical prote
low-density lipopr
low density lipopr
low density lipopr
hypothetical prote

C;Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein; 19/Domain: signal sequence #status predicted <SIG>
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-3943/Product: alpha-2-macroglobulin receptor <S15K>
 F;27-64/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F;72-108/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F;115-148/Domain: EGF homology <EG1>
 F;154-188/Domain: EGF homology <EG2>
 F;198-239/Domain: LDL receptor WYTD-containing repeat homology <YW01>
 F;240-381/Domain: LDL receptor WYTD-containing repeat homology <YW02>
 F;252-334/Domain: LDL receptor WYTD-containing repeat homology <YW03>
 F;335-378/Domain: LDL receptor WYTD-containing repeat homology <YW04>
 F;379-420/Domain: LDL receptor WYTD-containing repeat homology <YW05>
 F;421-468/Domain: LDL receptor WYTD-containing repeat homology <YW06>
 F;478-519/Domain: EGF homology <EG3>
 F;571-613/Domain: LDL receptor WYTD-containing repeat homology <YW07>
 F;614-659/Domain: LDL receptor WYTD-containing repeat homology <YW08>
 F;660-710/Domain: LDL receptor WYTD-containing repeat homology <YW09>
 F;711-752/Domain: LDL receptor WYTD-containing repeat homology <YW10>
 F;753-799/Domain: LDL receptor WYTD-containing repeat homology <YW11>
 F;807-842/Domain: EGF homology <EG4>
 F;854-890/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F;895-931/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F;936-971/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F;976-1011/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F;1015-1051/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F;1062-1097/Domain: LDL receptor ligand-binding repeat homology <LDL8>
 F;1104-1140/Domain: LDL receptor ligand-binding repeat homology <LDL9>
 F;1145-1182/Domain: LDL receptor ligand-binding repeat homology <LDL10>
 F;1185-1221/Domain: EGF homology <EG5>
 F;1227-1261/Domain: EGF homology <EG6>
 F;1269-1308/Domain: LDL receptor WYTD-containing repeat homology <YW12>
 F;1309-1358/Domain: LDL receptor WYTD-containing repeat homology <YW13>
 F;1356-1398/Domain: LDL receptor WYTD-containing repeat homology <YW14>
 F;1399-1445/Domain: LDL receptor WYTD-containing repeat homology <YW15>
 F;1446-1488/Domain: LDL receptor WYTD-containing repeat homology <YW16>
 F;1489-1531/Domain: LDL receptor WYTD-containing repeat homology <YW17>
 F;1540-1578/Domain: EGF homology <EG7>
 F;1583-1626/Domain: LDL receptor WYTD-containing repeat homology <YW18>
 F;1627-1669/Domain: LDL receptor WYTD-containing repeat homology <YW19>
 F;1670-1713/Domain: LDL receptor WYTD-containing repeat homology <YW20>
 F;1714-1753/Domain: LDL receptor WYTD-containing repeat homology <YW21>
 F;1754-1796/Domain: LDL receptor WYTD-containing repeat homology <YW22>
 F;1797-1846/Domain: LDL receptor WYTD-containing repeat homology <YW23>
 F;1850-1886/Domain: EGF homology <EG8>
 F;1934-1976/Domain: LDL receptor WYTD-containing repeat homology <YW24>
 F;1977-2019/Domain: LDL receptor WYTD-containing repeat homology <YW25>
 F;2020-2063/Domain: LDL receptor WYTD-containing repeat homology <YW26>
 F;2106-2151/Domain: LDL receptor WYTD-containing repeat homology <YW27>
 F;2159-2194/Domain: EGF homology <EG9>
 F;2199-2241/Domain: LDL receptor WYTD-containing repeat homology <YW28>
 F;2253-2294/Domain: LDL receptor WYTD-containing repeat homology <YW29>
 F;2344-2388/Domain: LDL receptor WYTD-containing repeat homology <YW30>
 F;2389-2429/Domain: LDL receptor WYTD-containing repeat homology <YW31>
 F;2430-2473/Domain: LDL receptor WYTD-containing repeat homology <YW32>
 F;2482-2517/Domain: EGF homology <EG10>
 F;2524-2561/Domain: LDL receptor ligand-binding repeat homology <LDL11>
 F;2566-2600/Domain: LDL receptor ligand-binding repeat homology <LDL12>
 F;2605-2639/Domain: LDL receptor ligand-binding repeat homology <LDL13>
 F;2652-2688/Domain: LDL receptor ligand-binding repeat homology <LDL14>
 F;2696-2730/Domain: LDL receptor ligand-binding repeat homology <LDL15>
 F;2734-2769/Domain: LDL receptor ligand-binding repeat homology <LDL16>
 F;2774-2812/Domain: LDL receptor ligand-binding repeat homology <LDL17>
 F;2818-2853/Domain: LDL receptor ligand-binding repeat homology <LDL18>
 F;2858-2897/Domain: LDL receptor ligand-binding repeat homology <LDL19>
 F;2904-2939/Domain: LDL receptor ligand-binding repeat homology <LDL20>
 F;2944-2980/Domain: EGF homology <EG11>
 F;2986-3021/Domain: EGF homology <EG12>
 F;3029-3068/Domain: LDL receptor WYTD-containing repeat homology <YW34>
 F;3069-3113/Domain: LDL receptor WYTD-containing repeat homology <YW35>
 F;3114-3156/Domain: LDL receptor WYTD-containing repeat homology <YW36>
 F;3157-3200/Domain: LDL receptor WYTD-containing repeat homology <YW37>
 F;3201-3241/Domain: LDL receptor WYTD-containing repeat homology <YW38>

F;3242-3284/Domain: LDL receptor WYTD-containing repeat homology <YW39>
 F;3294-3330/Domain: EGF homology <EG13>
 F;3334-3369/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F;3374-3408/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F;3413-3448/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F;3453-3488/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F;3494-3531/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F;3536-3570/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F;3574-3609/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F;3613-3647/Domain: LDL receptor ligand-binding repeat homology <LDL8>
 F;3654-3690/Domain: LDL receptor ligand-binding repeat homology <LDL9>
 F;3695-3731/Domain: LDL receptor ligand-binding repeat homology <LDL10>
 F;3741-3776/Domain: LDL receptor ligand-binding repeat homology <LDL11>
 F;3785-3822/Domain: EGF homology <EG14>
 F;3828-3860/Domain: EGF homology <EG15>
 F;3868-3911/Domain: LDL receptor WYTD-containing repeat homology <YW40>
 F;3912-3969/Domain: LDL receptor WYTD-containing repeat homology <YW41>
 F;3944-4544/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
 F;3944-4420/Domain: 85K chain extracellular #status predicted <EXT>
 F;3970-4012/Domain: LDL receptor WYTD-containing repeat homology <YW42>
 F;4013-4056/Domain: LDL receptor WYTD-containing repeat homology <YW43>
 F;4057-4099/Domain: LDL receptor WYTD-containing repeat homology <YW44>
 F;4100-4142/Domain: LDL receptor WYTD-containing repeat homology <YW45>
 F;4151-4182/Domain: EGF homology <EG16>
 F;4200-4231/Domain: EGF homology <EG17>
 F;4236-4267/Domain: EGF homology <EG18>
 F;4272-4303/Domain: EGF homology <EG19>
 F;4308-4339/Domain: EGF homology <EG20>
 F;4344-4374/Domain: EGF homology <EG21>
 F;4377-4408/Domain: EGF homology <EG22>
 F;4421-4444/Domain: transmembrane #status predicted <TMM>
 F;4445-4544/Domain: intracellular #status predicted <INT>
 F;166-2998/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
 F;2958/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
 F;4075,4125,4278/Binding site: carbohydrate (Asn) (covalent) #status predicted

Best Match 100.0%; Score 506; DB 1; Length 4544;
 Query Local Similarity 100.0%; Pred. No. 4.6e-38;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKXCFACEDQITCTISKWRCDCGERDCPDGSDAPEICQSKACRCQPNENCLGTEL 60
 Db 25 KTCSPKXCFACRQDQITCTISKWRCDCGERDCPDGSDAPEICQSKACRCQPNENCLGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
 Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 2
 S25111
 alpha-2-macroglobulin receptor precursor - mouse
 N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptor
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 22-Jun-1999
 C:Accession: S25111; S32554
 R:van Leuven, F.
 Submitted to the EMBL Data Library, July 1992
 A:Reference number: S25111
 A:Accession: S25111
 A:Molecule type: mRNA
 A:Residues: 1-4545 <VANI>
 A:Cross-references: EMBL:X67469; MID:g49941; PID:CAA47817.1; PID:g49942
 R:van Leuven, F.; Stas, L.; Raymakers, L.; Overbergh, L.; de Strooper, B.; Halliker, C.;
 Blochim. Biophys. Acta 1173, 71-74, 1993
 A:Title: Molecular cloning and sequencing of the murine alpha-2-macroglobulin receptor c
 A:Reference number: S32554; MUID:93250049; PMID:8485155
 A:Accession: S32554
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-28;4416-4453 <VAN2>
 A:Cross-references: EMBL:X67469
 C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated

ciated protein (see PIR:JX0281).

C/Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C/Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycopro
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-3944,3945-4545/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
F/20-3944/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
F/28-65/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/73-109/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/116-149/Domain: EGF homology <EG1>
F/155-189/Domain: EGF homology <EG2>
F/199-240/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F/241-282/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F/293-335/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F/336-379/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F/380-421/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F/422-469/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F/479-520/Domain: EGF homology <EG3>
F/522-614/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F/572-614/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F/615-660/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F/661-711/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F/712-753/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F/754-800/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F/808-843/Domain: EGF homology <EG4>
F/855-891/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F/896-932/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F/937-972/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F/977-1012/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F/1016-1052/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F/1063-1098/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F/1105-1141/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F/1146-1183/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F/1186-1222/Domain: EGF homology <EG5>
F/1228-1262/Domain: EGF homology <EG6>
F/1270-1309/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F/1310-1356/Domain: LDL receptor WYTD-containing repeat homology <YW13>
F/1357-1399/Domain: LDL receptor WYTD-containing repeat homology <YW14>
F/1400-1446/Domain: LDL receptor WYTD-containing repeat homology <YW15>
F/1447-1489/Domain: LDL receptor WYTD-containing repeat homology <YW16>
F/1490-1532/Domain: LDL receptor WYTD-containing repeat homology <YW17>
F/1541-1579/Domain: EGF homology <EG7>
F/1584-1627/Domain: LDL receptor WYTD-containing repeat homology <YW18>
F/1628-1670/Domain: LDL receptor WYTD-containing repeat homology <YW19>
F/1671-1714/Domain: LDL receptor WYTD-containing repeat homology <YW20>
F/1715-1754/Domain: LDL receptor WYTD-containing repeat homology <YW21>
F/1755-1797/Domain: LDL receptor WYTD-containing repeat homology <YW22>
F/1798-1847/Domain: LDL receptor WYTD-containing repeat homology <YW23>
F/1851-1887/Domain: EGF homology <EG8>
F/1935-1977/Domain: LDL receptor WYTD-containing repeat homology <YW24>
F/1978-2020/Domain: LDL receptor WYTD-containing repeat homology <YW25>
F/2021-2064/Domain: LDL receptor WYTD-containing repeat homology <YW26>
F/2065-2106/Domain: LDL receptor WYTD-containing repeat homology <YW27>
F/2107-2152/Domain: LDL receptor WYTD-containing repeat homology <YW28>
F/2160-2195/Domain: EGF homology <EG9>
F/2200-2242/Domain: LDL receptor WYTD-containing repeat homology <YW29>
F/2254-2295/Domain: LDL receptor WYTD-containing repeat homology <YW30>
F/2345-2389/Domain: LDL receptor WYTD-containing repeat homology <YW31>
F/2390-2436/Domain: LDL receptor WYTD-containing repeat homology <YW32>
F/2431-2474/Domain: LDL receptor WYTD-containing repeat homology <YW33>
F/2483-2518/Domain: EGF homology <EG10>
F/2525-2562/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F/2567-2601/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F/2606-2640/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F/2653-2689/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F/2697-2731/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F/2735-2770/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F/2775-2813/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F/2819-2854/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F/2859-2898/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F/2905-2940/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F/2945-2981/Domain: EGF homology <EG11>
F/2987-3022/Domain: LDL receptor WYTD-containing repeat homology <YW34>
F/3030-3069/Domain: LDL receptor WYTD-containing repeat homology <YW35>
F/3070-3114/Domain: LDL receptor WYTD-containing repeat homology <YW36>

F/3115-3157/Domain: LDL receptor WYTD-containing repeat homology <YW36>
F/3158-3201/Domain: LDL receptor WYTD-containing repeat homology <YW37>
F/3202-3242/Domain: LDL receptor WYTD-containing repeat homology <YW38>
F/3243-3285/Domain: LDL receptor WYTD-containing repeat homology <YW39>
F/3285-3331/Domain: EGF homology <EG13>
F/3335-3370/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/3375-3409/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/3414-3449/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F/3454-3490/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F/3495-3532/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F/3537-3571/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F/3576-3610/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F/3614-3648/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F/3655-3691/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F/3696-3732/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F/3742-3777/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F/3786-3823/Domain: EGF homology <EG14>
F/3829-3861/Domain: EGF homology <EG15>
F/3869-3912/Domain: LDL receptor WYTD-containing repeat homology <YW40>
F/3913-3970/Domain: LDL receptor WYTD-containing repeat homology <YW41>
F/3945-4545/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F/3945-4545/Domain: 85K chain extracellular #status predicted <EXT>
F/3971-4013/Domain: LDL receptor WYTD-containing repeat homology <YW42>
F/4014-4057/Domain: LDL receptor WYTD-containing repeat homology <YW43>
F/4058-4100/Domain: LDL receptor WYTD-containing repeat homology <YW44>
F/4101-4143/Domain: LDL receptor WYTD-containing repeat homology <YW45>
F/4152-4183/Domain: EGF homology <EG16>
F/4201-4232/Domain: EGF homology <EG17>
F/4237-4268/Domain: EGF homology <EG18>
F/4273-4304/Domain: EGF homology <EG19>
F/4309-4340/Domain: EGF homology <EG20>
F/4345-4375/Domain: EGF homology <EG21>
F/4378-4409/Domain: EGF homology <EG22>
F/4422-4445/Domain: transmembrane #status predicted <TM>
F/4446-4545/Domain: intracellular #status predicted <INT>
F/167,2999/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F/2959/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F/4076,4126,4279/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.0%; Score 486; DB 1; Length 4545;
Best Local Similarity 95.3%; Pred. No. 3e-36; Indels 0; Gaps 0;
Matches 82; Conservative 2; Mismatches 2

QY 1 KTCSPKPFACRQDQITCISKWRCDEGDCPDGSDGDEAFCICQSKAQRCPNHNCLGTEL 60
DB 26 KTCSPKPFACRQDQITCISKWRCDEGDCPDGSDGDEAFCICQSKAQRCPNHNCLGTEL 85

QY 61 CVPMSRLCNGVQDCWDGSDGEGPHORE 86
DB 86 CVPMSRLCNGVQDCWDGSDGEGPHORE 111

RESULT 3
A3102
Alpha-2-macroglobulin receptor precursor - chicken
N/Alternate names: C991; LDL receptor-related protein 1; low density lipoprotein receptor
C/Species: Gallus gallus (chicken)
C/Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 22-Jun-1999
C/Accession: A53102
R/Nimf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.
J. Biol. Chem. 269, 212-219, 1994
A/Title: The somatic cell-specific low density lipoprotein receptor-related protein of
C/Reference number: A53102; MUID:94103212; PMID:7506255
A/Accession: A53102
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-4543 <NIM>
A/Cross-references: GS:X74904; NID:G438006; PIDN:CAA52870.1; PID:G438007
C/Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated
d protein.
C/Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C/Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycopro
F/1-17/Domain: signal sequence #status predicted <SIG>

F:18-3942/Domain: alpha-2-macroglobulin receptor S15K chain #status predicted <S15K>
F:18-3942.3943/Domain: alpha-2-macroglobulin receptor #status predicted <MAT>
F:23-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:74-110/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:117-150/Domain: EGF homology <EG1>
F:156-190/Domain: EGF homology <EG2>
F:200-241/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F:242-283/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F:294-336/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F:337-380/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F:381-422/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F:423-470/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F:480-521/Domain: EGF homology <EG3>
F:573-615/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F:616-661/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F:662-712/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F:713-754/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:755-797/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:805-840/Domain: EGF homology <EG4>
F:852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:1183-1219/Domain: EGF homology <EG5>
F:1225-1259/Domain: EGF homology <EG6>
F:1267-1306/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:1307-1353/Domain: LDL receptor WYTD-containing repeat homology <YW13>
F:1354-1396/Domain: LDL receptor WYTD-containing repeat homology <YW14>
F:1444-1486/Domain: LDL receptor WYTD-containing repeat homology <YW15>
F:1487-1529/Domain: LDL receptor WYTD-containing repeat homology <YW16>
F:1538-1576/Domain: EGF homology <EG7>
F:1581-1624/Domain: LDL receptor WYTD-containing repeat homology <YW18>
F:1625-1667/Domain: LDL receptor WYTD-containing repeat homology <YW19>
F:1668-1711/Domain: LDL receptor WYTD-containing repeat homology <YW20>
F:1712-1751/Domain: LDL receptor WYTD-containing repeat homology <YW21>
F:1752-1794/Domain: LDL receptor WYTD-containing repeat homology <YW22>
F:1795-1842/Domain: LDL receptor WYTD-containing repeat homology <YW23>
F:1846-1882/Domain: EGF homology <EG8>
F:1930-1972/Domain: LDL receptor WYTD-containing repeat homology <YW24>
F:1973-2015/Domain: LDL receptor WYTD-containing repeat homology <YW25>
F:2016-2059/Domain: LDL receptor WYTD-containing repeat homology <YW26>
F:2060-2101/Domain: LDL receptor WYTD-containing repeat homology <YW27>
F:2102-2147/Domain: LDL receptor WYTD-containing repeat homology <YW28>
F:2155-2190/Domain: EGF homology <EG9>
F:2195-2237/Domain: LDL receptor WYTD-containing repeat homology <YW29>
F:2247-2288/Domain: LDL receptor WYTD-containing repeat homology <YW30>
F:2338-2382/Domain: LDL receptor WYTD-containing repeat homology <YW31>
F:2383-2423/Domain: LDL receptor WYTD-containing repeat homology <YW32>
F:2424-2467/Domain: LDL receptor WYTD-containing repeat homology <YW33>
F:2476-2511/Domain: EGF homology <EG10>
F:2518-2555/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F:2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F:2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F:2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F:2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F:2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F:2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F:2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDL20>
F:2941-2977/Domain: LDL receptor ligand-binding repeat homology <LDL21>
F:2983-3018/Domain: EGF homology <EG11>
F:3026-3065/Domain: LDL receptor WYTD-containing repeat homology <YW34>
F:3066-3110/Domain: LDL receptor WYTD-containing repeat homology <YW35>
F:3111-3153/Domain: LDL receptor WYTD-containing repeat homology <YW36>
F:3154-3197/Domain: LDL receptor WYTD-containing repeat homology <YW37>
F:3198-3238/Domain: LDL receptor WYTD-containing repeat homology <YW38>
F:3239-3281/Domain: LDL receptor WYTD-containing repeat homology <YW39>

F:3291-3327/Domain: EGF homology <EG13>
F:3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:3371-3405/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:3410-3445/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:3450-3486/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:3491-3528/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:3533-3567/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:3783-3820/Domain: EGF homology <EG14>
F:3826-3858/Domain: EGF homology <EG15>
F:3866-3909/Domain: LDL receptor WYTD-containing repeat homology <YW40>
F:3910-3968/Domain: LDL receptor WYTD-containing repeat homology <YW41>
F:3943-4011/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F:3943-4020/Domain: 85K chain extracellular #status predicted <EXT>
F:3959-4011/Domain: LDL receptor WYTD-containing repeat homology <YW42>
F:4012-4055/Domain: LDL receptor WYTD-containing repeat homology <YW43>
F:4056-4098/Domain: LDL receptor WYTD-containing repeat homology <YW44>
F:4099-4141/Domain: LDL receptor WYTD-containing repeat homology <YW45>
F:4150-4181/Domain: EGF homology <EG16>
F:4159-4230/Domain: EGF homology <EG17>
F:4235-4266/Domain: EGF homology <EG18>
F:4271-4302/Domain: EGF homology <EG19>
F:4307-4338/Domain: EGF homology <EG20>
F:4343-4373/Domain: EGF homology <EG21>
F:4376-4408/Domain: EGF homology <EG22>
F:4421-4443/Domain: transmembrane #status predicted <TM>
F:4444-4543/Domain: intracellular #status predicted <INT>
F:116,138,387,276,359,448,731,926,1048,1152,1153,1193,1216,1305,1509,1556,1573,1614,1643
3485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carboxydrate (Asn) (covalent)
F:168,2395/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

Query Match 90.7%; Score 459; DB 1; Length 4543;
Best Local Similarity 86.0%; Pred. No. 8 4e-34;
Matches 74; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KTCSPKQFACRQDITICISKWRCDGSRDPCDGSDEAPICPQSKAORCPNEHNCGLTEL 60
Db 27 KTCSPKQFACRQDITICISKWRCDGSRDPCDGSDEAPICPQSKAORCPNEHNCGLTEL 86

Qy 61 CVPMRLCNGVQDMGSDSGPHCRE 86
Db 87 CIEMSLCNGLHDFGSDSGPHCRE 112

RESULT 4
A47437
LDL-receptor-related protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 18-Aug-2000
C:Accession: A47437; S27801; T21547
R:Yochem, J.; Greenwald, I.
Proc. Natl. Acad. Sci. U.S.A. 90, 4572-4576, 1993
A:Title: A gene for a low density lipoprotein receptor-related protein in the nematode C.
A:Reference number: A47437; MUID:93281621; PMID:8506301
A:Accession: A47437
A:Molecule type: DNA
A:Residues: 1-4753 <YOC>
A:Cross-references: GB:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
A:Note: nucleotide sequence not given; translation not complete in this paper
R:Yochem, J.; Greenwald, I.
submitted to the EMBL Data Library, July 1992
A:Description: A gene for an LDL receptor-related protein (LPR) in the nematode C.elegans
A:Reference number: S27801
A:Accession: S27801
A:Molecule type: DNA
A:Residues: 1-4753 <YOC>
A:Cross-references: EMBL:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
R:Wilkinson, J.

RESULT 10

ORXLL1
LDL receptor 1 precursor - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Sep-1998
C/Accession: A40388
R/Moht, K.D.; Chen, W.J.; Goldstein, J.L.; Brown, M.S.
J. Biol. Chem. 266, 10406-10414, 1991
A/Title: The low density lipoprotein receptor in Xenopus laevis. Five domains that resemble the hamster low density lipoprotein receptor gene.
A/Reference number: A40388; MUID:91244815; PMID:1709931
A/Accession: A40388
A/Molecule type: mRNA
A/Residues: 1-909 <MEH>

A/Cross-references: GB:M62976
C/Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lipoprotein complex must first cluster into clathrin-coated pits.
C/Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology; LDL receptor; cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid
C/Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid
F/1-21/Domain: signal sequence #status predicted <SIG>
F/22-909/Product: LDL receptor #status predicted <MAT>
F/22-836/Domain: extracellular #status predicted <EXT>
F/27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F/148-183/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F/195-229/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F/234-268/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F/274-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F/316-350/Domain: EGF homology <EG1>
F/356-390/Domain: EGF homology <EG2>
F/397-436/Domain: LDL receptor WYTD-containing repeat homology <YW1>
F/437-483/Domain: LDL receptor WYTD-containing repeat homology <YW2>
F/484-526/Domain: LDL receptor WYTD-containing repeat homology <YW3>
F/527-570/Domain: LDL receptor WYTD-containing repeat homology <YW4>
F/571-613/Domain: LDL receptor WYTD-containing repeat homology <YW5>
F/614-656/Domain: LDL receptor WYTD-containing repeat homology <YW6>
F/665-709/Domain: EGF homology <EG3>
F/717-788/Region: clustered O-linked oligosaccharides
F/837-858/Domain: transmembrane #status predicted <TM>
F/859-909/Domain: intracellular #status predicted <INT>
F/873-877/Region: coated-pit mediated internalization signal
F/886-898/Region: basolateral targeting signal
F/97-270/459/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/316-327,323-336,338-350,356-366,362-375,377-390,665-679,675-694,696-709/Disulfide bond

Query Match 38.0%; Score 192.5; DB 1; Length 909;

Best Local Similarity 45.7%; Pred. No. 3.7e-10;

Matches 42; Conservative 9; Mismatches 30; Indels 11; Gaps 4;

QY 2 TCSPKQFACRDOITCISKWRCGERDCPDGSDPEAIC---PQSKAQRCPNHNCLG 57

Db 147 TCNPMFQCKDKGICIPKMWACDPPDCPDGSDPE--EHCEGRPIKTDKPCSPLEPHC-G 203

QY 58 TELCPVMSRLCNGVQDCMGSDSDE---GPHCR 85

Db 204 SGECHMSWKCDGDFCDKSKSDKDCVKPCTCR 235

RESULT 11

QRRLTD
LDL receptor precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C/Accession: S03430; S05076; B48908
R/Lee, L.Y.; Mohler, W.A.; Schaefer, B.L.; Freudenberger, J.S.; Byrne-Connolly, N.; Eager
Nucleic Acids Res. 17, 1259-1260, 1989
A/Title: Nucleotide sequence of the rat low density lipoprotein receptor cDNA.
A/Reference number: S03430; MUID:89160263; PMID:2922268
A/Accession: S03430
A/Molecule type: mRNA
A/Residues: 1-879 <LEE>
A/Cross-references: EMBL:X13722
R/Lee, L.

submitted to the EMBL Data Library, December 1988

A/Reference number: S05076

A/Accession: S05076

A/Molecule type: mRNA

A/Residues: 1-187,'D',189-879 <LE2>

A/Cross-references: EMBL:X13722; NID:956569; PIDN:CRAA32001.1; PID:956570

R/Bishop, R.W.

J. Lipid Res. 33, 549-557, 1992

A/Title: Structure of the hamster low density lipoprotein receptor gene.

A/Reference number: A48908; MUID:92407466; PMID:1527478

A/Accession: B48908

A/Molecule type: DNA

A/Residues: 1-7 <BIS>

C/Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lipoprotein complex must first cluster into clathrin-coated pits.

C/Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology; LDL receptor; cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid

C/Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid

F/1-21/Domain: signal sequence #status predicted <SIG>

F/22-879/Product: LDL receptor #status predicted <MAT>

F/22-807/Domain: extracellular #status predicted <EXT>

F/68-104/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/109-143/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F/148-184/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F/198-232/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F/237-271/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F/277-314/Domain: LDL receptor ligand-binding repeat homology <LDL6>

F/319-353/Domain: EGF homology <EG1>

F/359-393/Domain: EGF homology <EG2>

F/400-439/Domain: LDL receptor WYTD-containing repeat homology <YW1>

F/440-485/Domain: LDL receptor WYTD-containing repeat homology <YW2>

F/486-528/Domain: LDL receptor WYTD-containing repeat homology <YW3>

F/529-572/Domain: LDL receptor WYTD-containing repeat homology <YW4>

F/573-615/Domain: LDL receptor WYTD-containing repeat homology <YW5>

F/616-658/Domain: LDL receptor WYTD-containing repeat homology <YW6>

F/667-711/Domain: EGF homology <EG3>

F/717-788/Region: clustered O-linked oligosaccharides

F/808-829/Domain: transmembrane #status predicted <TM>

F/830-879/Domain: intracellular #status predicted <INT>

F/844-848/Region: coated-pit mediated internalization signal

F/856-868/Region: basolateral targeting signal

F/97-156,273,657/Binding site: carbohydrate (Asn) (covalent) #status predicted

F/319-330,326-339,341-353,359-369,365-378,380-393,667-681,677-696,698-711/Disulfide bond

F/717,720,721,724,726,732,733,748,757,766,768,775,780,785/Binding site: carbohydrate (T

F/725,734,738,746,755,764,767,769,779,783,787,788/Binding site: carbohydrate (Ser) (cov

Query Match 37.8%; Score 191.5; DB 1; Length 879;

Best Local Similarity 46.2%; Pred. No. 4.5e-10;

Matches 37; Conservative 10; Mismatches 28; Indels 5; Gaps 3;

QY 2 TCSPKQFACRDOITCISKWRCGERDCPDGSDPEAICPQSKAQRCPNHNCLG-TCL 60

Db 26 SCGNFEFCRDG-KCIYKWKVCDGRCPCDGSDESPETC---MSVTCRSGFSCGGRVSR 81

QY 61 CVPMSRLCNGVQDCMGSDSDE 80

Db 82 CIPDSWRCDCGRTDCNGSDE 101

RESULT 12

T29149
Hypothetical protein F47B3.8 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C/Accession: T29149
R/Du, Z.; Le, T.T.
submitted to the EMBL Data Library, April 1997
A/Description: The sequence of C. elegans cosmid F47B3.
A/Reference number: Z20579
A/Accession: T29149
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2643 <DUZ>

A;Cross-references: EMBL:U97017; PIDN:AA52363.1; GSPDB:GN00019; CESP:F47B3.8
 A;Experimental source: strain Bristol N2; clone F47B3

C;Genetics:

A;Gene: CESP:F47B3.8

A;Map position: 1

A;Introns: 29/1; 78/1; 124/1; 216/1; 256/1; 1222/3; 2036/3; 2127/2; 2478/3; 2522/3; 2589

C;Superfamily: LDL receptor ligand-binding repeat homology

F;44-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F;81-124/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F;130-166/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F;925-964/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F;969-1005/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F;1013-1047/Domain: LDL receptor ligand-binding repeat homology <LDL6>

F;1053-1087/Domain: LDL receptor ligand-binding repeat homology <LDL7>

F;1093-1132/Domain: LDL receptor ligand-binding repeat homology <LDL8>

F;1140-1175/Domain: LDL receptor ligand-binding repeat homology <LDL9>

F;1180-1216/Domain: LDL receptor ligand-binding repeat homology <LDL10>

F;1228-1265/Domain: LDL receptor ligand-binding repeat homology <LDL11>

Query Match 37.8%; Score 191.5; DB 2; Length 2643;

Best Local Similarity 46.4%; Pred. No. 1e-09;

Matches 39; Conservative 13; Mismatches 25; Indels 7; Gaps 5;

QY 3 CSPKQACRQDQITCISKWRCGERDCPDGSPDEAPICPSKAQRCOPNEHNCGLTCLV 62

DB 1013 CSVSQFQC-SKTKCIISKRCNGVQCDNGADE--EDCFRSKL--CDPDEFRC-GTGLCI 1066

QY 63 PMSRLCNGVQDCMDGSGPHCRE 86

DB 1067 KQSQVCDGKMQCLDGLDE-EHCNE 1089

RESULT 13

QRMSLD

LDL receptor precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1993 #sequence_revision 14-Aug-1998 #text_change 22-Jun-1999

A;Accession: I48623; JN0461

R;Polvino, W.J.; Dichek, D.A.; Mason, J.; Anderson, W.F.

Somat. Cell Mol. Genet. 18, 443-450, 1992

A;Title: Molecular cloning and nucleotide sequence of cDNA encoding a functional murine

A;Reference number: I48623; MUID:93117934; PMID:1475710

A;Accession: I48623

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-862 <RES>

A;Cross-references: EMBL:X64414; NID:G296201; PIDN:CAA45759.1; PID:G296202

R;Hoffer, M.J.V.; van Eck, M.W.; Petrif, F.; van der Zee, A.; de Wit, E.; Meijer, D.; GZ

Biochem. Biophys. Res. Commun. 191, 880-886, 1993

A;Title: The mouse low density lipoprotein receptor gene: cDNA sequence and exon-intron

A;Reference number: JN0461; MUID:93221531; PMID:8466528

A;Accession: JN0461

A;Molecule type: mRNA

A;Residues: 1-22 'V', 24-26 'G', 28-60 'K', 62-143 'P', 145-155 'K', 157-177 'H', 179-185 'AE'

A;Cross-references: GB:Z19521; NID:G296378; PIDN:CAA79581.1; PID:G296379

A;Experimental source: strain BAUB/C, adult male liver

C;Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lip

id complexes must first cluster into clathrin-coated pits.

C;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;

C;Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-862/Product: LDL receptor #status predicted <SIG>

F;22-790/Domain: LDL receptor #status predicted <EXT>

F;68-104/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F;109-143/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F;148-184/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F;198-232/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F;237-271/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F;277-314/Domain: LDL receptor ligand-binding repeat homology <LDL6>

F;319-353/Domain: LDL receptor ligand-binding repeat homology <EG1>

F;359-393/Domain: EGF homology <EG2>

F;400-439/Domain: LDL receptor YWTD-containing repeat homology <YW1>

F;440-485/Domain: LDL receptor YWTD-containing repeat homology <YW2>

F;486-528/Domain: LDL receptor YWTD-containing repeat homology <YW3>

F;529-572/Domain: LDL receptor YWTD-containing repeat homology <YW4>

F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW5>

F;616-658/Domain: LDL receptor YWTD-containing repeat homology <YW6>

F;667-712/Domain: EGF homology <EG3>

F;722-770/Region: clustered O-linked oligosaccharides

F;791-812/Domain: transmembrane #status predicted <TM>

F;813-862/Domain: intracellular #status predicted <INT>

F;827-831/Region: coated-pit mediated internalization signal

F;839-851/Region: basolateral targeting signal

F;97-273,452/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;319-330,326-339,341-353,359-369,365-378,380-393,667-682,678-697,699-712/Disulfide bond

Query Match 37.6%; Score 190.5; DB 1; Length 862;

Best Local Similarity 46.2%; Pred. No. 5.4e-10;

Matches 37; Conservative 11; Mismatches 27; Indels 5; Gaps 3;

QY 2 TCSPKQACRQDQITCISKWRCGERDCPDGSPDEAPICPSKAQRCOPNEHNCGLG-TEL 60

DB 26 SCSNRFQCRDG-KCTASKWVCDGSPCPDGSDESPETC---MSVTCQSNQFSGGRVSR 81

QY 61 CVPMSRLCNGVQDCMDGSGDE 80

DB 82 CIPDSWRCQVDCENDSDE 101

RESULT 14

QRMULD

LDL receptor precursor - Chinese hamster

C;Species: Cricetulus griseus (Chinese hamster)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999

A;Accession: A48908; A24426

R;Bishop, R.W.

J. Lipid Res. 33, 549-557, 1992

A;Title: Structure of the hamster low density lipoprotein receptor gene.

A;Reference number: A48908; MUID:92407468; PMID:1527478

A;Accession: A48908

A;Molecule type: DNA

A;Residues: 1-854 <BIS>

A;Cross-references: GB:U4387; NID:G191131; PIDN:AA51449.1; PID:G191132

A;Experimental source: UT-1 ovarian cell genomic library

A;Note: sequence extracted from NCBI backbone (NCBIN:113276, NCBIIP:113277)

R;Sege, R.D.; Kozarsky, K.F.; Krieger, M.

Mol. Cell. Biol. 6, 3268-3277, 1986

A;Title: Characterization of a family of gamma-ray-induced CHO mutants demonstrates that

A;Reference number: A24426; MUID:87064645; PMID:3785227

A;Accession: A24426

A;Molecule type: DNA

A;Residues: 570-615 <SEG>

C;Genetics:

A;Gene: ldlA

A;Introns: 23/1; 64/1; 105/1; 233/1; 274/1; 315/1; 355/1; 397/1; 454/2; 529/2; 569/1; 61

C;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;

C;Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-854/Product: LDL receptor #status predicted <EXT>

F;22-782/Domain: extracellular #status predicted <EXT>

F;27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F;68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F;109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F;148-184/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F;198-232/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F;237-271/Domain: LDL receptor ligand-binding repeat homology <LDL6>

F;277-314/Domain: LDL receptor ligand-binding repeat homology <LDL7>

F;319-353/Domain: EGF homology <EG1>

F;359-393/Domain: EGF homology <EG2>

F;400-439/Domain: LDL receptor YWTD-containing repeat homology <YW1>

F;440-485/Domain: LDL receptor YWTD-containing repeat homology <YW2>

F;486-528/Domain: LDL receptor YWTD-containing repeat homology <YW3>

F;529-572/Domain: LDL receptor YWTD-containing repeat homology <YW4>

F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW5>

F;616-658/Domain: LDL receptor YWTD-containing repeat homology <YW6>

Search completed: September 17, 2004, 11:18:37
Job time : 40 secs

```

Query Match      37.4%; Score 189; DB 1; Length 854;
Best Local Similarity 42.0%; Pred. No. 7.4e-10;
Matches 3; Conservative 15; Mismatches 22; Indels 14; Gaps 5;

Qy 2 TCSPKQACRQDIT-CISKGRCDGDRCDPGSDSE---APEICQSKAQRQCPNHNCLG 57
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 TCQKFEICGGRVSRCPNSNRCQGTDCENGSGDEQCAPKTCODEP-----RCDGK----- 120

```

RESULT 15
T29764
hypothetical protein T21E3.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T29764
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of *C. elegans* cosmid T21E3.
A:Reference number: Z20681
A:Accession: T29764
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2180 <DUZ>
A:Cross-references: EMBL:AF003133; PIDN:AAB54138.1; GSPDB:GN00019; CESP:T21E3.3
A:Experimental source: strain Bristol N2; clone T21E3
C:Genetics:
A:Gene: CESP:T21E3.3
A:Map position: 1
A:Introns: 400/2; 1608/3; 1674/3; 1686/2; 1731/2; 2073/1; 2098/2; 2146/

```

Query Match      37.4%; Score 189; DB 2; Length 2180;
Best Local Similarity 42.0%; Pred. No. 1.5e-09;
Matches 34; Conservative 15; Mismatches 26; Indels 6; Gaps 2;

QY 1 KTCSPK-QFACRDGITCISGWRCCGRDPCDGSDEAFETCPQSKAORCFQNEHNCIGTE 59
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 147 RSDAKTQFACLAIRTCMPKHWCCDGPDCADGSD-----KNCDAKCTCFEFCSCSSK 201
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:02:49 ; Search time 24 Seconds
(without alignments)
186.585 Million cell updates/sec

Title: US-09-625-137-21
Perfect score: 506
Sequence: 1 KTCSPKQFACRDQITCSKG.....LCNGVQDCMDGSDGPHCRE 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	506	100.0	4544 1 LRPI_HUMAN Q07954 homo sapien
2	459	90.7	4543 1 LRPI_CHICK P98157 gallus gall
3	204.5	40.4	4753 1 LRP_CAEEL Q04833 caenorhabdi
4	202	39.9	4660 1 LRP2_RAT P98158 rattus norv
5	197.5	39.0	1984 1 YL_DROME P98163 drosophila
6	192.5	38.0	909 1 LDLI_XENLA Q99087 xenopus lae
7	191.5	37.8	879 1 LDLR_RAT P35952 rattus norv
8	189	37.4	854 1 LDLR_CRIGR Q92473 h sortilin
9	187.5	37.1	2214 1 SORL_HUMAN Q06561 caenorhabdi
10	185	36.6	3375 1 UNS2_CAEEL P98156 mus musculu
11	184.5	36.5	873 1 LDVR_MOUSE Q95209 o sortilin
12	184.5	36.5	2213 1 SORL_RABIT P98164 homo sapien
13	183	36.2	4655 1 LRP2_HUMAN P20063 oryctolagus
14	182.5	36.1	837 1 LDLR_RABIT P98165 gallus gall
15	182.5	36.1	863 1 LDVR_CHICK P35953 oryctolagus
16	182.5	36.1	873 1 LDVR_RABIT Q98930 g sortilin
17	182.5	36.1	1592 1 SORL_CHICK P98166 rattus norv
18	181.5	35.9	873 1 LDVR_RAT Q99088 xenopus lae
19	181.5	35.9	892 1 LDLI_XENLA P98155 homo sapien
20	180.5	35.7	873 1 LDVR_MOUSE P35951 mus musculu
21	180	35.6	864 1 LDLI_MOUSE O88307 m sortilin
22	179.5	35.5	2215 1 SORL_MOUSE P01130 homo sapien
23	176	34.8	860 1 LDLR_HUMAN Q05793 mus musculu
24	174	34.4	4391 1 PGBM_HUMAN P98167 bos taurus
25	165	32.6	3707 1 PGBM_MOUSE Q95159 drosophila
26	159.5	31.5	2616 1 NDL_DROME Q92319 mus musculu
27	156.5	30.9	1113 1 CORI_MOUSE Q9Y5Y6 homo sapien
28	151.5	29.9	855 1 STI4_HUMAN P98167 bos taurus
29	151	29.8	867 1 SSPO_BOVIN P56677 mus musculu
30	141	27.9	855 1 STI4_MOUSE Q9Y5G5 homo sapien
31	140.5	27.8	1042 1 CORI_HUMAN P46023 lymnaea sca
32	139.5	27.6	1115 1 GPCE_LYMTST P34576 caenorhabdi
33	137	27.1	3767 1 MUA3_CAEEL

ALIGNMENTS

RESULT 1

ID	LRPI_HUMAN	STANDARD;	PRT;	4544 AA.
AC	Q07954;			
DT	01-OCT-1996	(Rel. 34, Created)		
DT	01-OCT-1996	(Rel. 34, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Low-density lipoprotein receptor-related protein 1 precursor (LRP)			
DE	(Alpha-2-macroglobulin receptor) (A2MR) (Apolipoprotein E receptor)			
DE	(APOER) (CD91).			
GN	LRP1 OR A2MR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=93210795; PubMed=3266596;			
RA	Herz J., Hamann U., Rognie S., Myklebost O., Gausepohl H.,			
RA	Stanley K.K.;			
RT	"Surface location and high affinity for calcium of a 500-kd liver			
RT	membrane protein closely related to the LDL-receptor suggest a			
RT	physiological role as lipoprotein receptor.";			
RL	EMBO J. 7:4119-4127(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95203893; PubMed=7534747;			
RA	Van Leuven F., Stas L., Hilliker C., Lorent K., Umans L., Serneels L.,			
RA	Overbergh L., Torrekens S., Moechars D., De Strooper B.,			
RA	Van den Bergh H.;			
RT	"Structure of the gene (LRP1) coding for the human alpha			
RT	2-macroglobulin receptor lipoprotein receptor-related protein.";			
RL	Genomics 24:78-89(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9900832; PubMed=9782078;			
RA	Van Leuven F., Stas L., Thiry E., Nelissen B., Miyake Y.;			
RT	"Strategy to sequence the 89 exons of the human LRP1 gene coding for			
RT	the lipoprotein receptor related protein: identification of one			
RT	expressed mutation among 48 polymorphisms.";			
RL	Genomics 52:138-144(1998).			
RN	[4]			
RP	PROCESSING.			
RX	MEDLINE=90269210; PubMed=2112085;			
RA	Herz J., Kowal R.C., Goldstein J.L., Brown M.S.;			
RT	"Proteolytic processing of the 600 kd low density lipoprotein			
RT	receptor-related protein (LRP) occurs in a trans-Golgi			
RT	compartment.";			
RL	EMBO J. 9:1769-1776(1990).			
RN	[5]			
RP	FUNCTION.			
RX	MEDLINE=91092405; PubMed=1702392;			
RA	Kristensen T., Moestrup S.K., Gliemann J., Bendtsen L., Sand O.,			
RA	Sottrup-Jensen L.;			
RT	"Evidence that the newly cloned low-density-lipoprotein receptor			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:02:49 ; Search time 24 Seconds
(without alignments)
186.585 Million cell updates/sec

Title: US-09-625-137-21
Perfect score: 506
Sequence: 1 KTCSPKQFACRDQITCSKG.....LCNGVQDCMDGSDGPHCRE 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	506	100.0	4544 1 LRPI_HUMAN Q07954 homo sapien
2	459	90.7	4543 1 LRPI_CHICK P98157 gallus gall
3	204.5	40.4	4753 1 LRP_CAEEL Q04833 caenorhabdi
4	202	39.9	4660 1 LRP2_RAT P98158 rattus norv
5	197.5	39.0	1984 1 YL_DROME P98163 drosophila
6	192.5	38.0	909 1 LDLI_XENLA Q99087 xenopus lae
7	191.5	37.8	879 1 LDLR_RAT P35952 rattus norv
8	189	37.4	854 1 LDLR_CRIGR Q92473 h sortilin
9	187.5	37.1	2214 1 SORL_HUMAN Q06561 caenorhabdi
10	185	36.6	3375 1 UNS2_CAEEL P98156 mus musculu
11	184.5	36.5	873 1 LDVR_MOUSE Q95209 o sortilin
12	184.5	36.5	2213 1 SORL_RABIT P98164 homo sapien
13	183	36.2	4655 1 LRP2_HUMAN P20063 oryctolagus
14	182.5	36.1	837 1 LDLR_RABIT P98165 gallus gall
15	182.5	36.1	863 1 LDVR_CHICK P35953 oryctolagus
16	182.5	36.1	873 1 LDVR_RABIT Q98930 g sortilin
17	182.5	36.1	1592 1 SORL_CHICK P98166 rattus norv
18	181.5	35.9	873 1 LDVR_RAT Q99088 xenopus lae
19	181.5	35.9	892 1 LDLI_XENLA P98155 homo sapien
20	180.5	35.7	873 1 LDVR_MOUSE P35951 mus musculu
21	180	35.6	864 1 LDLI_MOUSE O88307 m sortilin
22	179.5	35.5	2215 1 SORL_MOUSE P01130 homo sapien
23	176	34.8	860 1 LDLR_HUMAN Q05793 mus musculu
24	174	34.4	4391 1 PGBM_HUMAN P98167 bos taurus
25	165	32.6	3707 1 PGBM_MOUSE Q95159 drosophila
26	159.5	31.5	2616 1 NDL_DROME Q92319 mus musculu
27	156.5	30.9	1113 1 CORI_MOUSE Q9Y5Y6 homo sapien
28	151.5	29.9	855 1 STI4_HUMAN P98167 bos taurus
29	151	29.8	867 1 SSPO_BOVIN P56677 mus musculu
30	141	27.9	855 1 STI4_MOUSE Q9Y5G5 homo sapien
31	140.5	27.8	1042 1 CORI_HUMAN P46023 lymnaea sca
32	139.5	27.6	1115 1 GPCE_LYMTST P34576 caenorhabdi
33	137	27.1	3767 1 MUA3_CAEEL

related protein (LRP) is the alpha 2-macroglobulin receptor.";
PEBS Lett. 276:151-155(1990).
[6]
STRUCTURE BY NMR OF 1059-1100.
RP MEDLINE=99453972; PubMed=10318630;
RA Huang W., Dolmer K., Gettins P.G.W.;
RX "NMR solution structure of complement-like repeat CR8 from the low
RT density lipoprotein receptor-related protein.";
RN J. Biol. Chem. 274:14130-14136(1999).
[7]
STRUCTURE BY NMR OF 851-893.
RP MEDLINE=20119280; PubMed=10852313;
RA Dolmer K., Huang W., Gettins P.G.W.;
RX "NMR solution structure of complement-like repeat CR3 from the low
RT density lipoprotein receptor-related protein. Evidence for specific
RN binding to the receptor binding domain of human
alpha(2)-macroglobulin.";
RP J. Biol. Chem. 275:3284-3269(2000).
RL -1- FUNCTION: Involved in the plasma clearance of chylomicron remnants
CC and activated alpha 2-macroglobulin, as well as the local
CC metabolism of complexes between plasminogen activators and their
CC endogenous inhibitors.
CC
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC
CC -1- TISSUE SPECIFICITY: Most abundant in liver, brain and lung.
CC
CC -1- PM: Cleaved into a 85 kDa membrane-spanning subunit (LRP-85) and
CC a 515 kDa large extracellular domain (LRP-515) that remains non-
CC covalently associated.
CC
CC -1- SIMILARITY: Contains 22 EGF-like domains.
CC
CC -1- SIMILARITY: Contains 31 LDL-receptor class A domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13916; CAA32112.1; -;
CC EMBL; AF058427; AAC64265.1; -;
CC PIR; S02392; S02392.
CC PDB; 1CR8; 06-JAN-99.
CC PDB; 1D2L; 21-FEB-00.
CC PDB; 1J8E; 19-DEC-01.
CC Genew; HGNC:6694; LRP1.
CC MIM; 107770; -;
CC GO; GO:0005887; C: integral to plasma membrane; TAS.
CC GO; GO:0005624; C: membrane fraction; TAS.
CC GO; GO:0005509; F: calcium ion binding; TAS.
CC GO; GO:0005319; F: lipid transporter activity; TAS.
CC GO; GO:0008034; F: lipoprotein binding; TAS.
CC GO; GO:0004872; F: receptor activity; TAS.
CC GO; GO:0008283; F: cell proliferation; TAS.
CC GO; GO:0006629; P: lipid metabolism; TAS.
CC GO; GO:0009405; P: pathogenesis; TAS.
CC InterPro; IPR000152; Asx hydroxyl_s.
CC InterPro; IPR001881; EGF_CA.
CC InterPro; IPR008209; EGF_like.
CC InterPro; IPR002172; LDL_receptor.
CC InterPro; IPR000033; LDL_receptor_rep.
CC Pfam; PF00008; EGF; 14.
CC Pfam; PF00057; ldl_recept_a; 31.
CC Pfam; PF00058; ldl_recept_b; 33.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 3.
CC SMART; SM00192; LDLs; 31.
CC SMART; SM00135; Ly; 32.
CC PROSITE; PS00010; ASX_HYDROXYL; 3.
CC PROSITE; PS00022; EGF_1; 5.
CC PROSITE; PS01186; EGF_2; 8.
CC PROSITE; PS00026; EGF_3; 6.
CC PROSITE; PS01187; EGF_CA; 2.
CC PROSITE; PS00068; LDLRA_2; 31.
CC Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein; Signal;
CC Calcium-binding; EGF-like domain; Coated pits; 3D-structure;
CC Polymorphism.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 4544 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
CC DOMAIN 20 4419 PROTEIN 1.
CC TRANSMEM 4420 4444 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 4445 4544 POTENTIAL..
CC DOMAIN 25 66 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 70 110 LDL-RECEPTOR CLASS A 1.
CC DOMAIN 111 149 LDL-RECEPTOR CLASS A 2.
CC DOMAIN 150 189 EGF-LIKE 1.
CC DOMAIN 474 520 EGF-LIKE 2.
CC DOMAIN 803 843 EGF-LIKE 3.
CC DOMAIN 852 892 EGF-LIKE 4.
CC DOMAIN 893 933 LDL-RECEPTOR CLASS A 3.
CC DOMAIN 934 973 LDL-RECEPTOR CLASS A 4.
CC DOMAIN 974 1013 LDL-RECEPTOR CLASS A 5.
CC DOMAIN 1013 1053 LDL-RECEPTOR CLASS A 6.
CC DOMAIN 1060 1099 LDL-RECEPTOR CLASS A 7.
CC DOMAIN 1102 1142 LDL-RECEPTOR CLASS A 8.
CC DOMAIN 1143 1182 LDL-RECEPTOR CLASS A 9.
CC DOMAIN 1183 1222 EGF-LIKE 5.
CC DOMAIN 1223 1262 EGF-LIKE 6.
CC DOMAIN 1536 1579 EGF-LIKE 7.
CC DOMAIN 1846 1887 EGF-LIKE 8.
CC DOMAIN 2155 2195 EGF-LIKE 9.
CC DOMAIN 2478 2518 EGF-LIKE 10.
CC DOMAIN 2522 2563 LDL-RECEPTOR CLASS A 11.
CC DOMAIN 2564 2602 LDL-RECEPTOR CLASS A 12.
CC DOMAIN 2603 2641 LDL-RECEPTOR CLASS A 13.
CC DOMAIN 2642 2690 LDL-RECEPTOR CLASS A 14.
CC DOMAIN 2691 2732 LDL-RECEPTOR CLASS A 15.
CC DOMAIN 2732 2771 LDL-RECEPTOR CLASS A 16.
CC DOMAIN 2772 2814 LDL-RECEPTOR CLASS A 17.
CC DOMAIN 2816 2855 LDL-RECEPTOR CLASS A 18.
CC DOMAIN 2856 2899 LDL-RECEPTOR CLASS A 19.
CC DOMAIN 2902 2940 LDL-RECEPTOR CLASS A 20.
CC DOMAIN 2941 2981 EGF-LIKE 11.
CC DOMAIN 3022 3062 EGF-LIKE 12.
CC DOMAIN 3290 3331 EGF-LIKE 13.
CC DOMAIN 3332 3371 LDL-RECEPTOR CLASS A 21.
CC DOMAIN 3372 3410 LDL-RECEPTOR CLASS A 22.
CC DOMAIN 3411 3450 LDL-RECEPTOR CLASS A 23.
CC DOMAIN 3451 3491 LDL-RECEPTOR CLASS A 24.
CC DOMAIN 3492 3533 LDL-RECEPTOR CLASS A 25.
CC DOMAIN 3534 3572 LDL-RECEPTOR CLASS A 26.
CC DOMAIN 3573 3611 LDL-RECEPTOR CLASS A 27.
CC DOMAIN 3611 3649 LDL-RECEPTOR CLASS A 28.
CC DOMAIN 3652 3692 LDL-RECEPTOR CLASS A 29.
CC DOMAIN 3693 3733 LDL-RECEPTOR CLASS A 30.
CC DOMAIN 3739 3778 LDL-RECEPTOR CLASS A 31.
CC DOMAIN 3781 3823 EGF-LIKE 14.
CC DOMAIN 3824 3861 EGF-LIKE 15.
CC DOMAIN 4147 4183 EGF-LIKE 16.
CC DOMAIN 4184 4222 EGF-LIKE 17.
CC DOMAIN 4232 4268 EGF-LIKE 18.
CC DOMAIN 4268 4304 EGF-LIKE 19.
CC DOMAIN 4304 4340 EGF-LIKE 20.
CC DOMAIN 4340 4375 EGF-LIKE 21.
CC DOMAIN 4375 4409 EGF-LIKE 22.
CC DOMAIN 4409 4507 RECOGNITION SITE FOR PROTEOLYTICAL
CC SITE 3940 PROCESSING (POTENTIAL).
CC SITE 4502 ENDOCYTOSIS SIGNAL (POTENTIAL).
CC SITE 27 40 BY SIMILARITY.
CC DISULFID 34 53 BY SIMILARITY.
CC DISULFID 47 64 BY SIMILARITY.
CC DISULFID 72 85 BY SIMILARITY.
CC DISULFID 79 98 BY SIMILARITY.
CC DISULFID 92 108 BY SIMILARITY.
CC DISULFID 115 124 BY SIMILARITY.

Query Match 100.0%; Score 506; DB 1; Length 4544;
Best Local Similarity 100.0%; Pred. No. 6,2e-41; Indels 0; Gaps 0;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTSPKQFACRDIITCISKWRCGERDCPDGSDAEAPICPQSKAQCQFNEHNCIGTEL 60
Db 25 KTSPKQFACRDIITCISKWRCGERDCPDGSDAEAPICPQSKAQCQFNEHNCIGTEL 84
QY 61 CVPSRLCNGVQCMDSGEGPHCRE 86
Db 85 CVPSRLCNGVQCMDSGEGPHCRE 110

RESULT 2
LRP1_CHICK STANDARD; PRT; 4543 AA.

AC P98157; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Low-density lipoprotein receptor-related protein 1 precursor (LRP)
DE (Alpha-2-macroglobulin receptor) (A2MR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=White leghorn; TISSUE=Liver, and Ovary;
RX MEDLINE=94103212; PubMed=7506255;
RA Nimf J., Stifani S., Bilous P.T., Schneider W.J.;
RT "The somatic cell-specific low density lipoprotein receptor-related
protein of the chicken. Close kinship to mammalian low density
lipoprotein receptor gene family members.";
RT J. Biol. Chem. 269:212-219 (1994).
RL J. Biol. Chem. 269:212-219 (1994).
CC -!- FUNCTION: Involved in the plasma clearance of chylomicron remnants
and activated alpha 2-macroglobulin, as well as the local
metabolism of complexes of plasminogen activators and their
endogenous inhibitors. Binds vitellogenin, calcium and alpha 2-
macroglobulin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1;
CC IsoId=P98157-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P98157-2; Sequence=VSP_004312;
CC -!- TISSUE SPECIFICITY: Somatic.
CC -!- PTM: Cleaved into a 85 kDa membrane-spanning subunit (LRP-85) and
a 515 kDa large extracellular domain (LRP-515) that remains non-
covalently associated.
CC -!- SIMILARITY: Contains 22 EGF-like domains.
CC -!- SIMILARITY: Contains 31 LDL-receptor class A domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X74904; CAA52870.1; --
CC FDB; AS3102; A53102.
CC FDB; LRPX; 29-DEC-99.
CC InterPro; IPR000152; Asx hydroxyl_s.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00008; EGF; 14.

Pfam; PF00057; ldl_recept_a; 31.
Pfam; PF00058; ldl_recept_b; 33.
PRINTS; PRO0261; LDLRECEPTOR.
DR SMART; SMO0179; EGF_CA; 3.
DR SMART; SMO0192; LDLa; 31.
DR SMART; SMO0135; LY; 34.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 7.
DR PROSITE; PS00026; EGF_3; 8.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 31.
KW Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein; Signal;
KW Calcium-binding; EGF-like domain; Coated pits; Alternative splicing;
KW 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 4543
FT DOMAIN 22 4419
FT TRANSMEM 4420 4443
FT DOMAIN 4444 4543
FT DOMAIN 27 68
FT DOMAIN 72 112
FT DOMAIN 113 151
FT DOMAIN 152 191
FT DOMAIN 476 522
FT DOMAIN 801 841
FT DOMAIN 850 890
FT DOMAIN 891 931
FT DOMAIN 932 971
FT DOMAIN 972 1011
FT DOMAIN 1011 1051
FT DOMAIN 1058 1097
FT DOMAIN 1100 1140
FT DOMAIN 1141 1180
FT DOMAIN 1181 1220
FT DOMAIN 1221 1260
FT DOMAIN 1534 1577
FT DOMAIN 1842 1883
FT DOMAIN 2151 2191
FT DOMAIN 2472 2512
FT DOMAIN 2516 2557
FT DOMAIN 2558 2596
FT DOMAIN 2597 2635
FT DOMAIN 2636 2684
FT DOMAIN 2688 2730
FT DOMAIN 2730 2769
FT DOMAIN 2770 2812
FT DOMAIN 2814 2853
FT DOMAIN 2854 2897
FT DOMAIN 2900 2938
FT DOMAIN 2939 2978
FT DOMAIN 2979 3019
FT DOMAIN 3287 3328
FT DOMAIN 3329 3368
FT DOMAIN 3369 3407
FT DOMAIN 3408 3447
FT DOMAIN 3448 3488
FT DOMAIN 3489 3530
FT DOMAIN 3531 3569
FT DOMAIN 3570 3608
FT DOMAIN 3608 3646
FT DOMAIN 3649 3689
FT DOMAIN 3690 3730
FT DOMAIN 3736 3776
FT DOMAIN 3779 3821
FT DOMAIN 3822 3859
FT DOMAIN 4146 4182
FT DOMAIN 4195 4231
FT DOMAIN 4231 4267
FT DOMAIN 4267 4303
FT DOMAIN 4303 4339

POTENTIAL.
LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
PROTEIN 1.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
LDL-RECEPTOR CLASS A 1.
LDL-RECEPTOR CLASS A 2.
EGF-LIKE 1.
EGF-LIKE 2.
EGF-LIKE 3.
EGF-LIKE 4.
LDL-RECEPTOR CLASS A 3.
LDL-RECEPTOR CLASS A 4.
LDL-RECEPTOR CLASS A 5.
LDL-RECEPTOR CLASS A 6.
LDL-RECEPTOR CLASS A 7.
LDL-RECEPTOR CLASS A 8.
LDL-RECEPTOR CLASS A 9.
LDL-RECEPTOR CLASS A 10.
EGF-LIKE 5.
EGF-LIKE 6.
EGF-LIKE 7.
EGF-LIKE 8.
EGF-LIKE 9.
EGF-LIKE 10.
LDL-RECEPTOR CLASS A 11.
LDL-RECEPTOR CLASS A 12.
LDL-RECEPTOR CLASS A 13.
LDL-RECEPTOR CLASS A 14.
LDL-RECEPTOR CLASS A 15.
LDL-RECEPTOR CLASS A 16.
LDL-RECEPTOR CLASS A 17.
LDL-RECEPTOR CLASS A 18.
LDL-RECEPTOR CLASS A 19.
LDL-RECEPTOR CLASS A 20.
EGF-LIKE 11.
EGF-LIKE 12.
EGF-LIKE 13.
LDL-RECEPTOR CLASS A 21.
LDL-RECEPTOR CLASS A 22.
LDL-RECEPTOR CLASS A 23.
LDL-RECEPTOR CLASS A 24.
LDL-RECEPTOR CLASS A 25.
LDL-RECEPTOR CLASS A 26.
LDL-RECEPTOR CLASS A 27.
LDL-RECEPTOR CLASS A 28.
LDL-RECEPTOR CLASS A 29.
LDL-RECEPTOR CLASS A 30.
LDL-RECEPTOR CLASS A 31.
EGF-LIKE 14.
EGF-LIKE 15.
EGF-LIKE 16.
EGF-LIKE 17.
EGF-LIKE 18.
EGF-LIKE 19.
EGF-LIKE 20.

FT	DOMAIN	4339	4374	EGF-LIKE 21.
FT	DOMAIN	4372	4409	EGF-LIKE 22.
FT	SITE	3939	3942	RECOGNITION SITE FOR PROTEOLYTICAL
FT				PROCESSING (POTENTIAL).
FT				ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	SITE	4501	4506	BY SIMILARITY.
FT	DISULFID	29	42	BY SIMILARITY.
FT	DISULFID	36	55	BY SIMILARITY.
FT	DISULFID	49	66	BY SIMILARITY.
FT	DISULFID	74	87	BY SIMILARITY.
FT	DISULFID	81	100	BY SIMILARITY.
FT	DISULFID	94	110	BY SIMILARITY.
FT	DISULFID	117	126	BY SIMILARITY.
FT	DISULFID	122	135	BY SIMILARITY.
FT	DISULFID	137	150	BY SIMILARITY.
FT	DISULFID	156	166	BY SIMILARITY.
FT	DISULFID	162	175	BY SIMILARITY.
FT	DISULFID	177	190	BY SIMILARITY.
FT	DISULFID	480	495	BY SIMILARITY.
FT	DISULFID	491	506	BY SIMILARITY.
FT	DISULFID	508	521	BY SIMILARITY.
FT	DISULFID	805	816	BY SIMILARITY.
FT	DISULFID	812	825	BY SIMILARITY.
FT	DISULFID	827	840	BY SIMILARITY.
FT	DISULFID	852	864	BY SIMILARITY.
FT	DISULFID	859	877	BY SIMILARITY.
FT	DISULFID	871	888	BY SIMILARITY.
FT	DISULFID	893	905	BY SIMILARITY.
FT	DISULFID	900	918	BY SIMILARITY.
FT	DISULFID	912	929	BY SIMILARITY.
FT	DISULFID	934	946	BY SIMILARITY.
FT	DISULFID	941	959	BY SIMILARITY.
FT	DISULFID	953	969	BY SIMILARITY.
FT	DISULFID	974	987	BY SIMILARITY.
FT	DISULFID	982	1000	BY SIMILARITY.
FT	DISULFID	994	1009	BY SIMILARITY.
FT	DISULFID	1013	1025	BY SIMILARITY.
FT	DISULFID	1020	1038	BY SIMILARITY.
FT	DISULFID	1032	1049	BY SIMILARITY.
FT	DISULFID	1060	1073	BY SIMILARITY.
FT	DISULFID	1067	1086	BY SIMILARITY.
FT	DISULFID	1080	1095	BY SIMILARITY.
FT	DISULFID	1102	1116	BY SIMILARITY.
FT	DISULFID	1110	1129	BY SIMILARITY.
FT	DISULFID	1123	1138	BY SIMILARITY.
FT	DISULFID	1143	1157	BY SIMILARITY.
FT	DISULFID	1150	1170	BY SIMILARITY.
FT	DISULFID	1164	1180	BY SIMILARITY.
FT	DISULFID	1183	1194	BY SIMILARITY.
FT	DISULFID	1206	1204	BY SIMILARITY.
FT	DISULFID	1206	1219	BY SIMILARITY.
FT	DISULFID	1225	1235	BY SIMILARITY.
FT	DISULFID	1231	1244	BY SIMILARITY.
FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DISULFID	1258	1271	BY SIMILARITY.
FT	DISULFID	1288	1301	BY SIMILARITY.
FT	DISULFID	1301	1314	BY SIMILARITY.
FT	DISULFID	1314	1327	BY SIMILARITY.
FT	DISULFID	1327	1340	BY SIMILARITY.
FT	DISULFID	1340	1353	BY SIMILARITY.
FT	DISULFID	1353	1366	BY SIMILARITY.
FT	DISULFID	1366	1379	BY SIMILARITY.
FT	DISULFID	1379	1392	BY SIMILARITY.
FT	DISULFID	1392	1405	BY SIMILARITY.
FT	DISULFID	1405	1418	BY SIMILARITY.
FT	DISULFID	1418	1431	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1444	1457	BY SIMILARITY.
FT	DISULFID	1457	1470	BY SIMILARITY.
FT	DISULFID	1470	1483	BY SIMILARITY.
FT	DISULFID	1483	1496	BY SIMILARITY.
FT	DISULFID	1496	1509	BY SIMILARITY.
FT	DISULFID	1509	1522	BY SIMILARITY.
FT	DISULFID	1522	1535	BY SIMILARITY.
FT	DISULFID	1535	1548	BY SIMILARITY.
FT	DISULFID	1548	1561	BY SIMILARITY.
FT	DISULFID	1561	1574	BY SIMILARITY.
FT	DISULFID	1574	1587	BY SIMILARITY.
FT	DISULFID	1587	1600	BY SIMILARITY.
FT	DISULFID	1600	1613	BY SIMILARITY.
FT	DISULFID	1613	1626	BY SIMILARITY.
FT	DISULFID	1626	1639	BY SIMILARITY.
FT	DISULFID	1639	1652	BY SIMILARITY.
FT	DISULFID	1652	1665	BY SIMILARITY.
FT	DISULFID	1665	1678	BY SIMILARITY.
FT	DISULFID	1678	1691	BY SIMILARITY.
FT	DISULFID	1691	1704	BY SIMILARITY.
FT	DISULFID	1704	1717	BY SIMILARITY.
FT	DISULFID	1717	1730	BY SIMILARITY.
FT	DISULFID	1730	1743	BY SIMILARITY.
FT	DISULFID	1743	1756	BY SIMILARITY.
FT	DISULFID	1756	1769	BY SIMILARITY.
FT	DISULFID	1769	1782	BY SIMILARITY.
FT	DISULFID	1782	1795	BY SIMILARITY.
FT	DISULFID	1795	1808	BY SIMILARITY.
FT	DISULFID	1808	1821	BY SIMILARITY.
FT	DISULFID	1821	1834	BY SIMILARITY.
FT	DISULFID	1834	1847	BY SIMILARITY.
FT	DISULFID	1847	1860	BY SIMILARITY.
FT	DISULFID	1860	1873	BY SIMILARITY.
FT	DISULFID	1873	1886	BY SIMILARITY.
FT	DISULFID	1886	1899	BY SIMILARITY.
FT	DISULFID	1899	1912	BY SIMILARITY.
FT	DISULFID	1912	1925	BY SIMILARITY.
FT	DISULFID	1925	1938	BY SIMILARITY.
FT	DISULFID	1938	1951	BY SIMILARITY.
FT	DISULFID	1951	1964	BY SIMILARITY.
FT	DISULFID	1964	1977	BY SIMILARITY.
FT	DISULFID	1977	1990	BY SIMILARITY.
FT	DISULFID	1990	2003	BY SIMILARITY.
FT	DISULFID	2003	2016	BY SIMILARITY.
FT	DISULFID	2016	2029	BY SIMILARITY.
FT	DISULFID	2029	2042	BY SIMILARITY.
FT	DISULFID	2042	2055	BY SIMILARITY.
FT	DISULFID	2055	2068	BY SIMILARITY.
FT	DISULFID	2068	2081	BY SIMILARITY.
FT	DISULFID	2081	2094	BY SIMILARITY.
FT	DISULFID	2094	2107	BY SIMILARITY.
FT	DISULFID	2107	2120	BY SIMILARITY.
FT	DISULFID	2120	2133	BY SIMILARITY.
FT	DISULFID	2133	2146	BY SIMILARITY.
FT	DISULFID	2146	2159	BY SIMILARITY.
FT	DISULFID	2159	2172	BY SIMILARITY.
FT	DISULFID	2172	2185	BY SIMILARITY.
FT	DISULFID	2185	2198	BY SIMILARITY.
FT	DISULFID	2198	2211	BY SIMILARITY.
FT	DISULFID	2211	2224	BY SIMILARITY.
FT	DISULFID	2224	2237	BY SIMILARITY.
FT	DISULFID	2237	2250	BY SIMILARITY.
FT	DISULFID	2250	2263	BY SIMILARITY.
FT	DISULFID	2263	2276	BY SIMILARITY.
FT	DISULFID	2276	2289	BY SIMILARITY.
FT	DISULFID	2289	2302	BY SIMILARITY.
FT	DISULFID	2302	2315	BY SIMILARITY.
FT	DISULFID	2315	2328	BY SIMILARITY.
FT	DISULFID	2328	2341	BY SIMILARITY.
FT	DISULFID	2341	2354	BY SIMILARITY.
FT	DISULFID	2354	2367	BY SIMILARITY.
FT	DISULFID	2367	2380	BY SIMILARITY.
FT	DISULFID	2380	2393	BY SIMILARITY.
FT	DISULFID	2393	2406	BY SIMILARITY.
FT	DISULFID	2406	2419	BY SIMILARITY.
FT	DISULFID	2419	2432	BY SIMILARITY.
FT	DISULFID	2432	2445	BY SIMILARITY.
FT	DISULFID	2445	2458	BY SIMILARITY.
FT	DISULFID	2458	2471	BY SIMILARITY.
FT	DISULFID	2471	2484	BY SIMILARITY.
FT	DISULFID	2484	2497	BY SIMILARITY.
FT	DISULFID	2497	2510	BY SIMILARITY.
FT	DISULFID	2510	2523	BY SIMILARITY.
FT	DISULFID	2523	2536	BY SIMILARITY.
FT	DISULFID	2536	2549	BY SIMILARITY.
FT	DISULFID	2549	2562	BY SIMILARITY.
FT	DISULFID	2562	2575	BY SIMILARITY.
FT	DISULFID	2575	2588	BY SIMILARITY.

Matches	74;	Conservative	7;	Mismatches	5;	Indels	0;	Gaps	0;
Qy	1	KTCSPKQFACRQDITICISKGWCDGERDQPCDSDSAPEICQSKAQRCQPNHNCLGTEL	60						
Db	27	KTCSPKQFACRQDITICISKGWCDGERDQPCDSDSAPEICQSKAQRCQPNHNCLGTEL	86						
Qy	61	CVPMSELGVQDCMDGSDGSGHCRE	86						
Db	87	CIHMSKLCNGLRDCFDGSDGSGHCRE	112						
RESULT 3									
LRP_CAEEL STANDARD; PRT; 4753 AA.									
ID	Q04833;	01-OCT-1996 (Rel. 34, Created)							
DT	01-OCT-1996 (Rel. 34, Last sequence update)								
DT	28-FEB-2003 (Rel. 41, Last annotation update)								
DE	Low-density lipoprotein receptor-related protein precursor (LRP).								
GN	LRP-1 OR F29D11.1.								
OS	Caenorhabditis elegans.								
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;								
OC	Rhabditidae; Peloderinae; Caenorhabditis.								
OX	NCBI_TaxID=6239;								
RN	SEQUENCE FROM N.A.								
RP	MEDLINE=93281621; PubMed=8506301;								
RX	Yochem J., Greenwald I.								
RA	"A Gene for a low density lipoprotein receptor-related protein in the								
RL	nematode Caenorhabditis elegans."								
RL	Proc. Natl. Acad. Sci. U.S.A. 90:4572-4576 (1993).								
[2]									
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Bristol N2;								
RA	Wilkinson J.;								
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.								
CC	FUNCTION: May act as a receptor for the endocytosis of								
CC	extracellular ligands such as chylomicron remnants, protease-								
CC	inhibitor complexes and vitellogenin.								
CC	SUBCELLULAR LOCATION: Type I membrane protein.								
CC	SIMILARITY: Contains 35 LDL-receptor class A domains.								
CC	SIMILARITY: Contains 17 EGF-like domains.								
This SWISS-PROT entry is copyright. It is produced through a collaboration									
between the Swiss Institute of Bioinformatics and the EMBL outstation -									
the European Bioinformatics Institute. There are no restrictions on its									
use by non-profit institutions as long as its content is in no way									
modified and this statement is not removed. Usage by and for commercial									
entities requires a license agreement (See http://www.isb-sib.ch/announce/									
or send an email to license@isb-sib.ch).									
EMBL	M96150;	AAA28105.1;							
EMBL	Z73907;	CAA98124.1;							
PIR	A47437;	A47437.							
HSSP	Q07954;	1CR8.							
WormPep	F29D11.1;	CE05765.							
InterPro	IPR000152;	Asx_hydroxyl_5.							
InterPro	IPR001881;	EGF_Ca.							
InterPro	IPR006209;	EGF-like.							
InterPro	IPR002172;	LDL_receptor_A.							
InterPro	IPR000033;	ldl_receptor_rep.							
Pfam	PF00057;	ldl_recept_a;	34.						
Pfam	PF00058;	ldl_recept_b;	26.						
PRINTS	PR00261;	LDLRRECEPTOR.							
SMART	SM00179;	EGF_CA;	2.						
SMART	SM00192;	LDLr;	35.						
SMART	SM00135;	LY;	32.						
PROSITE	PS00010;	ASX_HYDROXYL;	6.						
PROSITE	PS00022;	EGF_1;	1.						
PROSITE	PS01186;	EGF_2;	3.						
PROSITE	PS50026;	EGF_3;	2.						
PROSITE	PS01187;	EGF_CA;	3.						
PROSITE	PS01209;	LDLRA_1;	27.						

[illegible]

Query Match 40.4%; Score 204.5; DB 1; Length 4753;
Best Local Similarity 36.0%; Pred. No. 7.1e-12;
Matches 41; Conservative 10; Mismatches 26; Indels 37;

3 CSPKQFACRDQITCISKWRC DGERDCPDGSD EAP EICP-----QSKAQRCQPNEH 53 QY

Db 182 CFQYQFRCADKTCIQKSWVCDGSKDCADGSD-PTDCEFFKKTANEFQCKNKRQPKRF 240
QY 54 -----NCLGTCLVPMRLNGVQDCWDGSDSE 80
Db 241 RCDYDDCGNDSDECGEYRCPGKWNCPGTHCHDIDQLKLDGSKOCADGADE 294
RESULT 4
LRP2_RAT
ID _LRP2_RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE (Glycoprotein 330) (gp330).
GN LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/megalin, a
RT distinctive member of the low density lipoprotein receptor gene
RT family";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RN [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gliemann J., Christensen E.I.;
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of
RT polybasic drugs";
RL J. Clin. Invest. 96:1404-1413(1995).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
RA Andres G., McCluskey R.T.;
RT "Organ distribution in rats of two members of the low-density
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the
RT receptor-associated protein (RAP).";
RL J. Histochem. Cytochem. 42:531-542(1994).
CC -1- FUNCTION: Binds plasminogen, extracellular matrix components,
CC plasminogen activator-plasminogen activator inhibitor type I
CC complex, apolipoprotein E-enriched beta-VLDL, lipoprotein lipase,
CC lactoferrin, clusterin and calcium.
CC -1- FUNCTION: Receptor-mediated uptake of polybasic drugs such as
CC aprotinin, aminoglycosides and polymyxin B.
CC -1- SUBUNIT: Forms a multimeric complex together with a receptor-
CC associated protein (RAP). Binds to ankyrin-repeat family A protein
CC 2 (ANKRA2) (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Expressed in
CC clathrin-coated pits; a soluble form is possibly derived by
CC cleavage at the cell surface.
CC -1- TISSUE SPECIFICITY: Epithelial cells of kidney glomerulus and
CC proximal tubule, lung, epididymis, yolk sac, among others.
CC -1- SIMILARITY: Contains 36 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 37 LDL-receptor class B domains.
CC -1- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sb.ch)

CC EMBL; L34049; AAA51369.1; --
DR PIR; T42737; T42737.
DR HSSP; Q07954; ICR8.
DR GlycoSuiteDB; P98158; --
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF00057; ldl_recept_a; 36.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLa; 36.
DR SMART; SM00135; LY; 34.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS00026; EGF_3; 8.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 31.
DR PROSITE; PS00068; LDLRA_2; 36.
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
KW Receptor; EGF-like domain; SH3-binding; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 4660 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT PROTEIN 2. EXTRACELLULAR (POTENTIAL).
FT TRANSMEM POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT LDL-RECEPTOR CLASS A 1.
FT LDL-RECEPTOR CLASS A 2.
FT LDL-RECEPTOR CLASS A 3.
FT LDL-RECEPTOR CLASS A 4.
FT LDL-RECEPTOR CLASS A 5.
FT LDL-RECEPTOR CLASS A 6.
FT LDL-RECEPTOR CLASS A 7.
FT EGF-LIKE 1.
FT EGF-LIKE 2.
FT LDL-RECEPTOR CLASS B 1.
FT LDL-RECEPTOR CLASS B 2.
FT LDL-RECEPTOR CLASS B 3.
FT LDL-RECEPTOR CLASS B 4.
FT LDL-RECEPTOR CLASS B 5.
FT EGF-LIKE 3.
FT LDL-RECEPTOR CLASS B 6.
FT LDL-RECEPTOR CLASS B 7.
FT LDL-RECEPTOR CLASS B 8.
FT LDL-RECEPTOR CLASS B 9.
FT EGF-LIKE 4.
FT LDL-RECEPTOR CLASS A 8.
FT LDL-RECEPTOR CLASS A 9.
FT LDL-RECEPTOR CLASS A 10.
FT LDL-RECEPTOR CLASS A 11.
FT LDL-RECEPTOR CLASS A 12.
FT LDL-RECEPTOR CLASS A 13.
FT LDL-RECEPTOR CLASS A 14.
FT LDL-RECEPTOR CLASS A 15.
FT EGF-LIKE 5.
FT EGF-LIKE 6. CALCIUM-BINDING (POTENTIAL).
FT LDL-RECEPTOR CLASS B 10.
FT LDL-RECEPTOR CLASS B 11.
FT LDL-RECEPTOR CLASS B 12.
FT LDL-RECEPTOR CLASS B 13.
FT EGF-LIKE 7.
FT LDL-RECEPTOR CLASS B 14.
FT LDL-RECEPTOR CLASS B 15.
FT LDL-RECEPTOR CLASS B 16.
FT LDL-RECEPTOR CLASS B 17.
FT LDL-RECEPTOR CLASS B 18.
FT LDL-RECEPTOR CLASS B 19.
FT LDL-RECEPTOR CLASS B 19.


```

CC EMBL; M62976; AAA49897.1; -.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR004209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00057; ldl_recept_a; 7.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SMO0179; EGF_CA; 1.
DR SMART; SMO0192; LDLa; 7.
DR SMART; SMO0135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; LDLA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 7.
KW Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 809 LOW-DENSITY LIPOPROTEIN RECEPTOR 1.
FT DOMAIN 22 836 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 837 858 POTENTIAL.
FT DOMAIN 859 909 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 65 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 66 106 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 107 145 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 146 185 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 193 231 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 232 270 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 272 311 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 312 351 EGF-LIKE 1.
FT DOMAIN 352 391 EGF-LIKE 2.
FT REPEAT 395 436 CALCIUM-BINDING (POTENTIAL).
FT REPEAT 437 483 LDL-RECEPTOR CLASS B 1.
FT REPEAT 484 526 LDL-RECEPTOR CLASS B 2.
FT REPEAT 527 570 LDL-RECEPTOR CLASS B 3.
FT REPEAT 571 613 LDL-RECEPTOR CLASS B 4.
FT REPEAT 614 655 LDL-RECEPTOR CLASS B 5.
FT DOMAIN 661 710 EGF-LIKE 3.
FT DOMAIN 717 813 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 871 876 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 27 39 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 68 82 BY SIMILARITY.
FT DISULFID 75 95 BY SIMILARITY.
FT DISULFID 89 104 BY SIMILARITY.
FT DISULFID 109 121 BY SIMILARITY.
FT DISULFID 116 134 BY SIMILARITY.
FT DISULFID 128 143 BY SIMILARITY.
FT DISULFID 148 161 BY SIMILARITY.
FT DISULFID 155 174 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 195 207 BY SIMILARITY.
FT DISULFID 202 220 BY SIMILARITY.
FT DISULFID 214 229 BY SIMILARITY.
FT DISULFID 234 246 BY SIMILARITY.
FT DISULFID 241 259 BY SIMILARITY.
FT DISULFID 253 268 BY SIMILARITY.
FT DISULFID 274 287 BY SIMILARITY.
FT DISULFID 282 300 BY SIMILARITY.
FT DISULFID 294 311 BY SIMILARITY.
FT DISULFID 316 327 BY SIMILARITY.
FT DISULFID 323 336 BY SIMILARITY.
FT DISULFID 338 350 BY SIMILARITY.
FT DISULFID 356 366 BY SIMILARITY.

```

```

FT DISULFID 362 375 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 665 679 BY SIMILARITY.
FT DISULFID 675 694 BY SIMILARITY.
FT DISULFID 696 709 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 909 AA; 101295 MW; 6ED41F5402A16371 CRC64;

Query Match 38.0%; Score 192.5; DB 1; Length 909;
Best Local Similarity 45.7%; Pred. No. 2.3e-11;
Matches 42; Conservative 9; Mismatches 30; Indels 11; Gaps 4;

QY 2 TCSPKQAFACRDQITCISKGRDGDGSDGDEPAEIC----POSKAQRQCPNEHNCIG 57
Db 147 TCNPAMFOCKDKGICIPKLWACDGDGDCEDGSDSDE--EHCEGREPIKTKPCSPLEFHC-G 203
QY 58 TELCVPMRLNGVODCMGDSDE----GPHCR 85
Db 204 SGECHMSWKDGGDFDKDKSDKDCVKTCTCR 235

RESULT 7
LDLR RAT
ID LDLR RAT STANDARD; PRT; 879 AA.
AC P33952;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Low-density lipoprotein receptor precursor (LDL receptor).
GN LDLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:Sprague-Dawley;
RX MEDLINE=89160263; PubMed=2922268;
RA Lee L.Y., Mohler W.A., Schafer B.L., Freudenberger J.S.,
RA Byrne-Connolly N., Eager K.B., Mosley S.T., Leighton J.K.,
RA Thrift R.N., Davis R.A., Tanaka R.D.;
RT "Nucleotide sequence of the rat low density lipoprotein receptor
RT cDNA."
RL Nucleic Acids Res. 17:1259-1260(1989).
CC -!- FUNCTION: Binds LDL, the major cholesterol-carrying lipoprotein of
CC plasma, and transports it into cells by endocytosis. In order to
CC be internalized, the receptor-ligand complexes must first cluster
CC into clathrin-coated pits.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13722; CAA32001.1; -.
DR PIR; S03430; QRRULD.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 7.

```


FT	DOMAIN	1197	LDL-RECEPTOR CLASS A 1.
FT	DOMAIN	1237	LDL-RECEPTOR CLASS A 5.
FT	DOMAIN	1273	LDL-RECEPTOR CLASS A 6.
FT	DOMAIN	1317	LDL-RECEPTOR CLASS A 7.
FT	DOMAIN	1323	LDL-RECEPTOR CLASS A 8.
FT	DOMAIN	1366	LDL-RECEPTOR CLASS A 8.
FT	DOMAIN	1417	LDL-RECEPTOR CLASS A 9.
FT	DOMAIN	1469	LDL-RECEPTOR CLASS A 10.
FT	DOMAIN	1512	LDL-RECEPTOR CLASS A 11.
FT	DOMAIN	1556	FIBRONECTIN TYPE-III 1.
FT	DOMAIN	1653	FIBRONECTIN TYPE-III 2.
FT	DOMAIN	1749	FIBRONECTIN TYPE-III 3.
FT	DOMAIN	1842	FIBRONECTIN TYPE-III 4.
FT	DOMAIN	1933	FIBRONECTIN TYPE-III 5.
FT	DOMAIN	2025	FIBRONECTIN TYPE-III 6.
FT	SITE	63	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	2172	ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	DISULFID	1078	BY SIMILARITY.

```

SEQUENCE      224 aa;          Score 197.5; DB 1; Length 2214;
Query Match    37.1%;          Pred. No. 1.6e-10; Indels 11;
Best Local Similarity 47.6%; Mismatches 24;
Matches 40; Conservative 9;
      QY   2 TCSPQFACRDIITCISGWRCDEGDCPDSRAPEICPOS-KAQRCQPNEHNCV
            |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||
Db        1198 TCEASNFQCRNG-HCIPQRWACDGTDCQSGSDEDPVNCERKCKGFC-PN-----
      QY   61 CYPMSLGNVGQDCWDGSDGEPHC 84
            |||.:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db        1249 CIPSSKHCDGLRDCSDGSDE-QHC 1271

RESULT 10
UN52 CAEEL
ID _UN52 CAEEL STANDARD; PRT; 3375 AA.
AC Q06561; O18263; O18263; Q9XTD2; Q9XTI5;
DT 01-JUN-1994 (Rel. 29, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane proteoglycan precursor (Perlecan homolog)
DE Uncoordinated protein 52).

```

GN UNC-52 OR ZC101.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.
RX MEDLINE=9339574; PubMed=939416;
RA Regalski T.M., Williams B.D., Mullen G.P., Moerman D.G.;
RT "Products of the unc-52 gene in Caenorhabditis elegans are homologous
RT to the core protein of the mammalian basement membrane heparan
RT sulfate proteoglycan."
RL Genes Dev. 7:1471-1484 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Percy C.M., Baynes C.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Durbin R.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Probable role in myofibrillar assembly and/or attachment
CC of the myofibrillar lattice to the cell membrane. May be an
CC extracellular anchor for integrin receptors in muscle.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=a;
CC IsoId=Q06561-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=a;
CC IsoId=Q06561-2; Sequence=VSP_007195, VSP_007196;
CC Name=b;
CC IsoId=Q06561-3; Sequence=VSP_007191, VSP_007192;
CC Note=No experimental confirmation available;
CC Name=c;
CC IsoId=Q06561-4; Sequence=VSP_007193, VSP_007194, VSP_007195,
CC VSP_007196;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Found in the basement membrane of all
CC contractile tissues. It is concentrated over muscle dense bodies
CC and M-lines which are associated with beta-integrin.
CC -!- DEVELOPMENTAL STAGE: Synthesized early in embryogenesis.
CC -!- SIMILARITY: Contains 3 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 7 laminin G-like domains.
CC -!- SIMILARITY: Contains 3 laminin G-like domains.
CC -!- SIMILARITY: Contains 2 laminin IV domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13458; AAA28156.1; -
CC EMBL; Z93375; CAB07567.1; -
CC EMBL; Z93395; CAB07567.1; JOINED.
CC EMBL; Z93375; CAB07568.1; -
CC EMBL; Z93395; CAB07568.1; JOINED.
CC EMBL; Z93375; CAB07569.1; -
CC EMBL; Z93395; CAB07569.1; JOINED.
CC EMBL; Z93395; CAB07569.1; JOINED.
CC EMBL; Z93395; CAB07704.1; -
CC EMBL; Z93395; CAB07706.1; -
CC EMBL; Z93375; CAB07706.1; JOINED.
CC EMBL; Z93395; CAB07707.1; -
CC EMBL; Z93375; CAB07707.1; JOINED.
CC EMBL; Z93395; CAB07708.1; -
CC EMBL; Z93375; CAB07708.1; JOINED.
CC HSSP; P01130; 1LDR.

DR WormPep; ZC101.2a; CE15028.
DR WormPep; ZC101.2b; CE15030.
DR WormPep; ZC101.2c; CE15034.
DR WormPep; ZC101.2e; CE18424.
DR GO; GO:0005578; C:extracellular matrix; IEPP.
DR GO; GO:0030239; P:myofibril assembly; IEPP.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF-like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000034; Laminin B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00047; Ig; 16.
DR Pfam; PF00052; laminin_B; 2.
DR Pfam; PF00053; laminin_EGF; 5.
DR Pfam; PF00057; ldl_recept_a; 3.
DR PRINTS; PRO0261; LDLRECEPTOR.
DR ProDom; PD003031; Laminin_B; 2.
DR SMART; SM0018; EGF; 6.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00180; EGF_Lam; 6.
DR SMART; SM00409; IG; 17.
DR SMART; SM00408; IGC2; 17.
DR SMART; SM00281; LamB; 1.
DR SMART; SM00282; LamG; 3.
DR SMART; SM00192; LDLA; 3.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 3.
DR PROSITE; PS00835; IG_Like; 17.
DR PROSITE; PS00025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 7.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
KW Proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;
KW Basement membrane; Extracellular matrix; Alternative splicing;
KW Laminin EGF-like domain.
FT SIGNAL 1 22
FT CHAIN 23 3375
FT DOMAIN 45 130
FT DOMAIN 148 184
FT DOMAIN 189 225
FT DOMAIN 232 269
FT DOMAIN 271 355
FT DOMAIN 384 431
FT DOMAIN 432 441
FT DOMAIN 442 633
FT DOMAIN 634 666
FT DOMAIN 674 720
FT DOMAIN 721 730
FT DOMAIN 731 921
FT DOMAIN 922 954
FT DOMAIN 955 1004
FT DOMAIN 1011 1060
FT DOMAIN 1061 1111
FT DOMAIN 1126 1222
FT DOMAIN 1226 1311
FT DOMAIN 1319 1401
FT DOMAIN 1410 1499
FT DOMAIN 1503 1585
FT DOMAIN 1588 1680
FT DOMAIN 1690 1785
FT DOMAIN 1793 1878
FT DOMAIN 1886 1970
FT DOMAIN 1973 2069
FT DOMAIN 2073 2163
FT DOMAIN 2173 2260
FT DOMAIN 2263 2343
BASMENT MEMBRANE PROTEOGLYCAN.
IG-LIKE C2-TYPE 1.
LDL-RECEPTOR CLASS A 1.
LDL-RECEPTOR CLASS A 2.
LDL-RECEPTOR CLASS A 3.
IG-LIKE C2-TYPE 2.
LAMININ EGF-LIKE 1 (INCOMPLETE).
LAMININ EGF-LIKE 2 (N-TERMINAL).
LAMININ DOMAIN IV 1.
LAMININ EGF-LIKE 2 (C-TERMINAL).
LAMININ EGF-LIKE 3 (INCOMPLETE).
LAMININ EGF-LIKE 4 (N-TERMINAL).
LAMININ DOMAIN IV 2.
LAMININ EGF-LIKE 4 (C-TERMINAL).
LAMININ EGF-LIKE 5.
LAMININ EGF-LIKE 6.
LAMININ EGF-LIKE 7.
IG-LIKE C2-TYPE 3.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 5.
IG-LIKE C2-TYPE 6.
IG-LIKE C2-TYPE 7.
IG-LIKE C2-TYPE 8.
IG-LIKE C2-TYPE 9.
IG-LIKE C2-TYPE 10.
IG-LIKE C2-TYPE 11.
IG-LIKE C2-TYPE 12.
IG-LIKE C2-TYPE 13.
IG-LIKE C2-TYPE 14.
IG-LIKE C2-TYPE 15.

FT	DOMAIN	2349	2435	IG-LIKE C2-TYPE 16.	OX	NCBI_TaxID=10090;
FT	DOMAIN	2446	2530	IG-LIKE C2-TYPE 17.	RN	[1]
FT	DOMAIN	2532	2713	LAMININ G-LIKE 1.	RP	SEQUENCE FROM N.A.
FT	DOMAIN	2793	2960	LAMININ G-LIKE 2.	RC	STRAIN=BALB/c; TISSUE=Heart;
FT	DOMAIN	2961	3093	GLU-RICH.	RX	MEDLINE=95010090; PubMed=7925422;
FT	DOMAIN	2972	3066	THR-RICH.	RA	Oka K., Ishimura-Oka K., Chu M.J., Sullivan M., Krushkal J.,
FT	DOMAIN	3180	3359	LAMININ G-LIKE 3.	RA	Li W.H., Chan L.;
FT	DISULFID	66	114	BY SIMILARITY.	RA	"Mouse very-low-density-lipoprotein receptor (VLDLR) cDNA cloning,
FT	DISULFID	149	161	BY SIMILARITY.	RT	tissue-specific expression and evolutionary relationship with the
FT	DISULFID	156	174	BY SIMILARITY.	RT	low-density-lipoprotein receptor.";
FT	DISULFID	168	183	BY SIMILARITY.	RL	Eur. J. Biochem. 224:975-982(1994).
FT	DISULFID	190	202	BY SIMILARITY.	RN	[2]
FT	DISULFID	197	215	BY SIMILARITY.	RP	SEQUENCE FROM N.A.
FT	DISULFID	209	224	BY SIMILARITY.	RC	TISSUE=Skeletal muscle;
FT	DISULFID	233	246	BY SIMILARITY.	RX	MEDLINE=94283285; PubMed=8013374;
FT	DISULFID	240	259	BY SIMILARITY.	RA	Gafvels M.E., Paavola L.G., Boyd C.O., Nolan P.M., Wittmaack F.,
FT	DISULFID	253	268	BY SIMILARITY.	RA	Chawla A., Lazar M.A., Bucan M., Angelin B.O., Strauss J.F.;
FT	DISULFID	955	964	BY SIMILARITY.	RT	"Cloning of a complementary deoxyribonucleic acid encoding the murine
FT	DISULFID	957	971	BY SIMILARITY.	RT	homolog of the very low density lipoprotein/apolipoprotein-E
FT	DISULFID	974	983	BY SIMILARITY.	RT	receptor: expression pattern and assignment of the gene to mouse
FT	DISULFID	986	1002	BY SIMILARITY.	RT	chromosome 19.";
FT	DISULFID	1011	1021	BY SIMILARITY.	RT	Endocrinology 135:387-394 (1994).
FT	DISULFID	1013	1027	BY SIMILARITY.	RN	[3]
FT	DISULFID	1030	1039	BY SIMILARITY.	RP	SEQUENCE OF 204-262 FROM N.A.
FT	DISULFID	1042	1058	BY SIMILARITY.	RX	MEDLINE=95003355; PubMed=7919660;
FT	DISULFID	1061	1069	BY SIMILARITY.	RA	Neggert J.K., Wu J.L.;
FT	DISULFID	1063	1079	BY SIMILARITY.	RA	"The mouse very low density lipoprotein receptor (Vldlr) gene maps to
FT	DISULFID	1082	1091	BY SIMILARITY.	RT	chromosome 19.";
FT	DISULFID	1094	1109	BY SIMILARITY.	RL	Mamm. Genome 5:453-455(1994).
FT	DISULFID	1152	1200	BY SIMILARITY.	RN	[4]
FT	DISULFID	1338	1384	BY SIMILARITY.	RP	BINDING TO REELIN.
FT	DISULFID	1435	1481	BY SIMILARITY.	RX	MEDLINE=20036019; PubMed=10571241;
FT	DISULFID	1527	1573	BY SIMILARITY.	RA	Hiesberger T., Tromsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
FT	DISULFID	1618	1663	BY SIMILARITY.	RA	Cooper J.A., Herz J.;
FT	DISULFID	1719	1767	BY SIMILARITY.	RT	"Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
FT	DISULFID	1814	1861	BY SIMILARITY.	RT	tyrosine phosphorylation of disabled-1 and modulates tau
FT	DISULFID	1907	1954	BY SIMILARITY.	RT	phosphorylation.";
FT	DISULFID	1998	2053	BY SIMILARITY.	RL	Neuron 24:481-489(1999).
FT	DISULFID	2099	2147	BY SIMILARITY.	CC	-1- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
FT	DISULFID	2195	2242	BY SIMILARITY.	CC	In order to be internalized, the receptor-ligand complexes must
FT	DISULFID	2284	2329	BY SIMILARITY.	CC	first cluster into clathrin-coated pits. Binding to Reelin induces
FT	DISULFID	2374	2420	BY SIMILARITY.	CC	tyrosine phosphorylation of Dab1 and modulation of tau
FT	DISULFID	2467	2514	BY SIMILARITY.	CC	phosphorylation.
FT	CARBOHYD	1422	1422	N-LINKED (GLCNAC. .) (POTENTIAL).	CC	-1- SUBUNIT: Binds to the extracellular matrix protein Reelin.
FT	CARBOHYD	2476	2476	N-LINKED (GLCNAC. .) (POTENTIAL).	CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
FT	CARBOHYD	2950	2950	N-LINKED (GLCNAC. .) (POTENTIAL).	CC	-1- TISSUE SPECIFICITY: Abundant in heart and muscle; less in kidney,
FT	CARBOHYD	3143	3143	N-LINKED (GLCNAC. .) (POTENTIAL).	CC	brain, ovary, testis, lung and adipose tissue.
Qy	1	KTSPKQFACDQITCISKWRCGDCGDCGSDPEICPQSKACQPCPNEHNCIGTEL 60			CC	-1- SIMILARITY: Contains 8 LDL-receptor class A domains.
Db	188	RTCPNEFEKNNN-KCVQKWLCDGDDCGDNDDEL-NCNAKPSDDCKPTBFQCHDRRQ 245			CC	-1- SIMILARITY: Contains 3 EGF-like domains.
Qy	61	CVPMRLCNGVQDCMDGSD 80			CC	-----
Db	246	CVSSFFHCGTNDCHDGSDE 265			CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
RESULT 11					CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
ID	LDVR_MOUSE	STANDARD;	PRT;	873 AA.	CC	the European Bioinformatics Institute. There are no restrictions on its
AC	P98156; 064022;				CC	use by non-profit institutions as long as its content is in no way
DT	01-OCT-1996 (Rel. 34, Created)				CC	modified and this statement is not removed. Usage by and for commercial
DT	01-OCT-1996 (Rel. 34, Last sequence update)				CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
DT	15-MAR-2004 (Rel. 43, Last annotation update)				CC	or send an email to license@sib-sib.ch).
DE	Very low-density lipoprotein receptor precursor (VLDL receptor).				CC	-----
GN	VLDLR.				CC	EMBL; L33417; AAC37668.1; -
OS	Mus musculus (Mouse).				DR	EMBL; U06670; AAA59384.1; -
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				DR	EMBL; S73732; AAB32228.2; -
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				DR	PIR; I48952; I48952.
					DR	HSSP; P01130; 1A3J.
					DR	MGD; MGI:98935; Vldlr.
					DR	InterPro; IPR000152; Asx_hydroxyl_s.
					DR	InterPro; IPR001881; EGF_Ca.
					DR	InterPro; IPR006209; EGF_like.
					DR	InterPro; IPR002172; LDL_receptor_A.
					DR	InterPro; IPR000033; Ldl_receptor_rep.
					DR	Pfam; PF00008; EGF; 2.
					DR	Pfam; PF00057; ldl_recept_a; 8.
					DR	Pfam; PF00058; ldl_recept_b; 5.

RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,
RA Rask L., Morse B.;
RT "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor
with potential intracellular signaling properties.";
RL Eur. J. Biochem. 239:132-137 (1996).
RN [2]
RP SEQUENCE OF 2705-4453 FROM N.A.
RC TISSUE=Kidney;
RA Knaak C., Agraves W.S.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 3933-4453 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95048397; PubMed=7959795;
RA Korenberg J.R., Agraves K.M., Chen X.N., Tran H.,
RA Strickland D.K., Agraves W.S.;
RT "Chromosomal localization of human genes for the LDL receptor family
member glycoprotein 330 (LRP2) and its associated protein RAP
(LRPAP1).";
RT Genomics 22:88-93 (1994).
RN [4]
RP SEQUENCE OF 4139-4406 FROM N.A.
RX MEDLINE=94244704; PubMed=8187828;
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,
RA Klarskog L., Akerstrom G., Rask L.;
RT "A protein involved in calcium sensing of the human parathyroid and
placental cytotrophoblast cells belongs to the LDL-receptor protein
superfamily.";
RL Exp. Cell Res. 212:344-350 (1994).
RN [5]
RP FUNCTION
RX MEDLINE=95286588; PubMed=7769501;
RA Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,
RA Brewer B.H., Strickland D.K., Agraves W.S.;
RT "Identification of glycoprotein 330 as an endocytic receptor for
apolipoprotein J/clusterin.";
RL J. Biol. Chem. 270:13070-13075 (1995).
CC -1- FUNCTION: Binds specifically clusterin with high affinity, but
also ligands in common with other family members: plasminogen,
extracellular matrix components, plasminogen activator-plasminogen
activator inhibitor type I complex, apolipoprotein E-enriched
beta-VLDL, lipoprotein lipase, lactoferrin and calcium.
CC -1- FUNCTION: receptor-mediated uptake of polybasic drugs such as
apronin, aminoglycosides and polymyxin B (By similarity).
CC -1- FUNCTION: May participate in regulation of parathyroid-hormone and
para-thyroid-hormone-related protein release.
CC -1- SUBUNIT: Forms a multimeric complex together with a receptor-
associated protein (RAP). Binds to ankyrin-repeat family A protein
2 (ANKRA2) (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Absorptive epithelia, including renal
proximal tubules.
CC -1- SIMILARITY: Contains 36 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 37 LDL-receptor class B domains.
CC -1- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33837; AB41649.1; --
DR EMBL; U04441; AB02882.1; --
DR EMBL; S73145; AB30825.1; --
DR PIR; I53413; I53413.
DR HSSP; Q07954; 1CR8.
DR Genew; HGNC:6694; LRP2.
DR MIM; 600073; --
GO; GO:0005764; C:lysosome; TAS.

GO; GO:0006629; P:lipid metabolism; TAS.
GO; GO:0006486; P:protein amino acid glycosylation; TAS.
GO; GO:0006898; P:receptor mediated endocytosis; TAS.
DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00057; ldl_recept_a; 36.
DR Pfam; PF00058; ldl_recept_b; 37.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLA; 36.
DR SMART; SM00135; LY; 33.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS00026; EGF_3; 6.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 31.
DR PROSITE; PS00068; LDLRA_2; 36.
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
KW Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 4555
FT DOMAIN 26 4423
FT TRANSMEM 4424 4446
FT DOMAIN 4447 4555
FT DOMAIN 26 64
FT DOMAIN 65 105
FT DOMAIN 106 144
FT DOMAIN 145 181
FT DOMAIN 182 219
FT DOMAIN 220 258
FT DOMAIN 264 308
FT DOMAIN 309 347
FT DOMAIN 348 386
FT DOMAIN 436 477
FT DOMAIN 479 520
FT DOMAIN 522 567
FT DOMAIN 569 612
FT DOMAIN 613 653
FT DOMAIN 659 705
FT DOMAIN 753 794
FT DOMAIN 796 836
FT DOMAIN 838 880
FT DOMAIN 882 924
FT DOMAIN 970 1014
FT DOMAIN 1024 1062
FT DOMAIN 1065 1103
FT DOMAIN 1107 1145
FT DOMAIN 1147 1185
FT DOMAIN 1186 1224
FT DOMAIN 1228 1268
FT DOMAIN 1269 1307
FT DOMAIN 1310 1350
FT DOMAIN 1349 1389
FT DOMAIN 1390 1429
FT DOMAIN 1478 1519
FT DOMAIN 1521 1562
FT DOMAIN 1566 1608
FT DOMAIN 1610 1653
FT DOMAIN 1655 1695
FT DOMAIN 1700 1741
FT DOMAIN 1790 1831
FT DOMAIN 1833 1881
FT DOMAIN 1883 1929
FT DOMAIN 1931 1971
FT DOMAIN 1972 2012
FT DOMAIN 2018 2059
FT DOMAIN 2107 2155
EXTRACELLULAR (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
LDL-RECEPTOR CLASS A 1.
LDL-RECEPTOR CLASS A 2.
LDL-RECEPTOR CLASS A 3.
LDL-RECEPTOR CLASS A 4.
LDL-RECEPTOR CLASS A 5.
LDL-RECEPTOR CLASS A 6.
LDL-RECEPTOR CLASS A 7.
EGF-LIKE 1.
EGF-LIKE 2.
LDL-RECEPTOR CLASS B 1.
LDL-RECEPTOR CLASS B 2.
LDL-RECEPTOR CLASS B 3.
LDL-RECEPTOR CLASS B 4.
LDL-RECEPTOR CLASS B 5.
EGF-LIKE 3.
LDL-RECEPTOR CLASS B 6.
LDL-RECEPTOR CLASS B 7.
LDL-RECEPTOR CLASS B 8.
LDL-RECEPTOR CLASS B 9.
EGF-LIKE 4.
LDL-RECEPTOR CLASS A 8.
LDL-RECEPTOR CLASS A 9.
LDL-RECEPTOR CLASS A 10.
LDL-RECEPTOR CLASS A 11.
LDL-RECEPTOR CLASS A 12.
LDL-RECEPTOR CLASS A 13.
LDL-RECEPTOR CLASS A 14.
LDL-RECEPTOR CLASS A 15.
EGF-LIKE 5.
EGF-LIKE 6.
CALCIUM-BINDING (POTENTIAL).
LDL-RECEPTOR CLASS B 10.
LDL-RECEPTOR CLASS B 11.
LDL-RECEPTOR CLASS B 12.
LDL-RECEPTOR CLASS B 13.
LDL-RECEPTOR CLASS B 14.
EGF-LIKE 7.
LDL-RECEPTOR CLASS B 15.
LDL-RECEPTOR CLASS B 16.
LDL-RECEPTOR CLASS B 17.
LDL-RECEPTOR CLASS B 18.
LDL-RECEPTOR CLASS B 19.
EGF-LIKE 8.
LDL-RECEPTOR CLASS B 20.

```

FT DOMAIN 2157 2200 LDL-RECEPTOR CLASS B 21.
FT DOMAIN 2202 2244 LDL-RECEPTOR CLASS B 22.
FT DOMAIN 2246 2288 LDL-RECEPTOR CLASS B 23.
FT DOMAIN 2290 2331 LDL-RECEPTOR CLASS B 24.
FT DOMAIN 2342 2383 EGF-LIKE 9.
FT DOMAIN 2431 2476 LDL-RECEPTOR CLASS B 25.
FT DOMAIN 2478 2517 LDL-RECEPTOR CLASS B 26.
FT DOMAIN 2519 2561 LDL-RECEPTOR CLASS B 27.
FT DOMAIN 2563 2603 LDL-RECEPTOR CLASS B 28.
FT DOMAIN 2604 2645 LDL-RECEPTOR CLASS B 29.
FT DOMAIN 2651 2693 EGF-LIKE 10.
FT DOMAIN 2698 2738 LDL-RECEPTOR CLASS A 16.
FT DOMAIN 2739 2777 LDL-RECEPTOR CLASS A 17.
FT DOMAIN 2778 2819 LDL-RECEPTOR CLASS A 18.
FT DOMAIN 2820 2861 LDL-RECEPTOR CLASS A 19.
FT DOMAIN 2862 2901 LDL-RECEPTOR CLASS A 20.
FT DOMAIN 2904 2945 LDL-RECEPTOR CLASS A 21.
FT DOMAIN 2946 2990 LDL-RECEPTOR CLASS A 22.
FT DOMAIN 2991 3029 LDL-RECEPTOR CLASS A 23.
FT DOMAIN 3030 3070 LDL-RECEPTOR CLASS A 24.
FT DOMAIN 3073 3110 LDL-RECEPTOR CLASS A 25.
FT DOMAIN 3111 3151 EGF-LIKE 11.
FT DOMAIN 3152 3192 EGF-LIKE 12.
FT DOMAIN 3193 3230 LDL-RECEPTOR CLASS B 30.
FT DOMAIN 3231 3268 LDL-RECEPTOR CLASS B 31.
FT DOMAIN 3269 3307 LDL-RECEPTOR CLASS B 32.
FT DOMAIN 3308 3346 LDL-RECEPTOR CLASS B 33.
FT DOMAIN 3347 3385 LDL-RECEPTOR CLASS B 34.
FT DOMAIN 3386 3424 EGF-LIKE 13.
FT DOMAIN 3425 3463 LDL-RECEPTOR CLASS A 26.
FT DOMAIN 3464 3502 LDL-RECEPTOR CLASS A 27.
FT DOMAIN 3503 3541 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3542 3580 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3581 3619 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3620 3658 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3659 3697 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3698 3736 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3737 3775 LDL-RECEPTOR CLASS A 34.
FT DOMAIN 3776 3814 LDL-RECEPTOR CLASS A 35.
FT DOMAIN 3815 3853 LDL-RECEPTOR CLASS A 36.

Query Match 36.2%; Score 183; DB 1; Length 4655;
Best Local Similarity 42.9%; Pred. No. 8.2e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRDIQTICISKWRCDGDCPDGSDPEAIPCQAKQCPQNEHNCGLGTEL 60
Db 1269 KTCPSYFHC-DNGNCIHRALWCDRNDGCDMSDE--KDCP-TQPRCFSQWQCLGHNI 1324

QY 61 CVPMSRLCNVQDCMGSDGPHC 84
Db 1325 CYNLSVVCDFGFCPNGTDSPLC 1348

RESULT 14
LDLR_RABIT STANDARD; PRT; 837 AA.
AC P20063;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Low-density lipoprotein receptor (LDL receptor) (Fragment).
GN LDLR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86208167; PubMed=3010466;
RA Yamamoto T., Bishop R.W., Brown M.S., Goldstein J.L., Russell D.W.;
RT "Deletion in cysteine-rich region of LDL receptor impedes transport
RT to cell surface in WHRL rabbit."
RL Science 232:1230-1237(1986).
CC -!- FUNCTION: Binds LDL, the major cholesterol-carrying lipoprotein of

```

```

CC plasma, and transports it into cells by endocytosis. In order to
CC be internalized, the receptor-ligand complexes must first cluster
CC into clathrin-coated pits.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11501; AAA31383.1; -.
CC PIR; A29512; A29512.
CC HSP; P01130; LDLR.
CC InterPro; IPR000152; Asx_hydroxyl_S.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00057; ldl_recept_a; 7.
CC Pfam; PF00058; ldl_recept_b; 5.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00192; LDLa; 7.
CC SMART; SM00135; LY; 5.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 2.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01209; LDLRA_1; 7.
CC PROSITE; PS00068; LDLRA_2; 7.
CC Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; EGF-like domain;
KW Repeat.
FT DOMAIN 1 765 EXTRACELLULAR (POTENTIAL).
FT DOMAIN <1 765 POTENTIAL.
FT TRANSMEM 766 787 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 788 837 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 12 52 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 53 93 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 94 132 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 133 173 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 182 220 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 221 259 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 261 300 EGF-LIKE 1.
FT DOMAIN 301 340 EGF-LIKE 2.
FT DOMAIN 341 380 EGF-LIKE 3.
FT REPEAT 384 425 LDL-RECEPTOR CLASS B 1.
FT REPEAT 426 471 LDL-RECEPTOR CLASS B 2.
FT REPEAT 473 515 LDL-RECEPTOR CLASS B 3.
FT REPEAT 516 559 LDL-RECEPTOR CLASS B 4.
FT REPEAT 560 602 LDL-RECEPTOR CLASS B 5.
FT REPEAT 603 644 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 650 699 EGF-LIKE 3.
FT SITE 800 805 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 14 26 BY SIMILARITY.
FT DISULFID 21 39 BY SIMILARITY.
FT DISULFID 33 59 BY SIMILARITY.
FT DISULFID 55 69 BY SIMILARITY.
FT DISULFID 62 82 BY SIMILARITY.
FT DISULFID 76 91 BY SIMILARITY.
FT DISULFID 96 108 BY SIMILARITY.
FT DISULFID 103 121 BY SIMILARITY.
FT DISULFID 115 130 BY SIMILARITY.
FT DISULFID 135 147 BY SIMILARITY.
FT DISULFID 142 160 BY SIMILARITY.

```


FT	DISULFID	152	167	BY SIMILARITY.
FT	DISULFID	172	184	BY SIMILARITY.
FT	DISULFID	179	197	BY SIMILARITY.
FT	DISULFID	191	206	BY SIMILARITY.
FT	DISULFID	211	223	BY SIMILARITY.
FT	DISULFID	218	236	BY SIMILARITY.
FT	DISULFID	230	247	BY SIMILARITY.
FT	DISULFID	257	269	BY SIMILARITY.
FT	DISULFID	264	282	BY SIMILARITY.
FT	DISULFID	276	291	BY SIMILARITY.
FT	DISULFID	296	308	BY SIMILARITY.
FT	DISULFID	303	321	BY SIMILARITY.
FT	DISULFID	315	330	BY SIMILARITY.
FT	DISULFID	336	349	BY SIMILARITY.
FT	DISULFID	344	362	BY SIMILARITY.
FT	DISULFID	356	373	BY SIMILARITY.
FT	DISULFID	378	389	BY SIMILARITY.
FT	DISULFID	385	398	BY SIMILARITY.
FT	DISULFID	400	412	BY SIMILARITY.
FT	DISULFID	418	428	BY SIMILARITY.
FT	DISULFID	424	437	BY SIMILARITY.
FT	DISULFID	439	452	BY SIMILARITY.
FT	DISULFID	726	739	BY SIMILARITY.
FT	DISULFID	735	754	BY SIMILARITY.
FT	DISULFID	756	769	BY SIMILARITY.
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	773	773	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	863 AA;	94904 MW;	0672A8748F9A2245 CRC64;
Query Match 36.1%; Score 182.5; DB 1; Length 863;				
Best Local Similarity 46.2%; Pred. No. 2e-10;				
Matches 36; Conservative 11; Mismatches 24; Indels 7; Gaps 4;				
QY	3	CSPKQACRQDITCISKGWRCDCGERDCPDGSDDEAPICPQSKAQCOPNEHNCGLGTELCV	62	
Db	257	CSTSEVQCGSG-ECIHKKWRCDDPCKDGSDEIN--CP---SRTCRPDQFRCEDGN-CI	309	
QY	63	PMRLCNGVQDCWDGSDSDE	80	
Db	310	HGSRQCNGVRDCLDGTDE	327	

Search completed: September 17, 2004, 11:15:51
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:09:54 ; Search time 116 Seconds
(without alignments)
233.919 Million cell updates/sec

Title: US-09-625-137-21
Perfect score: 506
Sequence: 1 KTCSPKQFACRDOITCISK.....LCNGVDCMGSDGEGPHCRE 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	292	4 Q86SW0	Q86sw0 homo sapien
2	506	100.0	296	4 Q727K9	Q727k9 homo sapien
3	486	96.0	4545	11 Q912X7	Q912x7 mus musculu
4	486	96.0	4545	11 Q920Y4	Q920y4 mus musculu
5	486	96.0	4545	11 Q61291	Q61291 mus musculu
6	280	55.3	881	4 Q8WY31	Q8wy31 homo sapien
7	280	55.3	4599	4 Q9NZR2	Q9nzt2 homo sapien
8	270	53.4	4599	11 Q9J118	Q9ji18 mus musculu
9	209	41.3	883	5 Q46131	Q46131 locusta nig
10	207.5	41.0	883	5 Q9VBN1	Q9vbn1 drosophila
11	205.5	40.6	4699	5 Q9V383	Q9v383 drosophila
12	204.5	40.4	891	5 Q7YW57	Q7yw57 aedes aegypt
13	204.5	40.4	1156	5 Q963T3	Q963t3 aedes aegypt
14	203	40.1	1537	4 Q8WY29	Q8wy29 homo sapien
15	202	39.9	739	5 Q8IGR9	Q8igr9 drosophila
16	202	39.9	826	5 Q8GB77	Q8gb77 drosophila

17	202	39.9	861	5 Q7YT26	Q7yt26 drosophila
18	202	39.9	1064	5 Q7YU01	Q7yu01 drosophila
19	202	39.9	1069	5 Q9VBN2	Q9vbn2 drosophila
20	197.5	38.0	134	5 Q95QH2	Q95qh2 caenorhabdi
21	197.5	39.0	1937	5 Q9VY56	Q9vy56 drosophila
22	197.5	39.0	1937	5 Q86P52	Q86p52 drosophila
23	196.5	38.8	355	13 Q802Y2	Q802y2 brachydanio
24	194	38.3	820	4 Q96NT6	Q96nt6 homo sapien
25	193	38.1	870	11 Q921B6	Q921b6 mus musculu
26	193	38.1	951	4 Q86V27	Q86v27 homo sapien
27	193	38.1	963	4 Q14114	Q14114 homo sapien
28	192.5	38.0	847	13 Q90W12	Q90w12 oncorhynch
29	192.5	38.0	891	13 Q7T2X3	Q7t2x3 gallus gall
30	192	37.9	844	13 Q7ZTG7	Q7ztg7 oreochromis
31	191.5	37.8	379	13 Q7SKY0	Q7sky0 brachydanio
32	191.5	37.8	699	4 Q99876	Q99876 homo sapien
33	191.5	37.8	793	4 Q14968	Q14968 homo sapien
34	191.5	37.8	2643	5 Q01552	Q01552 caenorhabdi
35	190.5	37.6	862	11 Q912J1	Q912j1 mus musculu
36	190.5	37.4	2192	5 Q01768	Q01768 caenorhabdi
37	189	37.4	1650	11 Q9QV76	Q9qv76 rattus sp.
38	187.5	37.1	519	11 Q8CCP5	Q8ccp5 mus musculu
39	187	37.0	996	11 Q924X6	Q924x6 mus musculu
40	187	37.0	1905	11 Q8V156	Q8v156 mus musculu
41	187	37.0	870	6 Q02660	Q02660 bos taurus
42	186.5	36.9	5146	6 Q8SPM4	Q8spm4 bos taurus
43	185.5	36.7	869	13 Q42126	Q42126 xenopus lae
44	185.5	36.7	1081	5 Q8T4N8	Q8t4n8 penaeus sem

ALIGNMENTS

RESULT 1
Q86SW0 PRELIMINARY; PRT; 292 AA.
ID Q86SW0;
AC Q86SW0;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to low density lipoprotein-related protein 1
DE (Alpha-2-macroglobulin receptor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045107; AA45107.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00057; LDL_recept_a; 2.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLa; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 2.
KW Lipoprotein.
SQ SEQUENCE 292 AA; 31631 MW; 82C39315F1BCE24 CRC64;

Wed Sep 22 12:33:00 2004

us-09-625-137-21.rspt

```

Query Match      100.0%; Score 506; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 5.5e-55;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 84

QY 61 CVPMSRLCNGVQDCWMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCWMDGSDGPHCRE 110

RESULT 2
Q727K9 PRELIMINARY; PRT; 296 AA.
AC Q727K9;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.O., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052593; AAH52593.1; -.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 32226 MW; C3FEF9F2F13E5854 CRC64;

Query Match      100.0%; Score 506; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 5.6e-55;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 84

QY 61 CVPMSRLCNGVQDCWMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCWMDGSDGPHCRE 110

RESULT 3
Q91ZX7 PRELIMINARY; PRT; 4545 AA.
AC Q91ZX7;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA;
RA Roebroek A.J.M.;
RT "LRP cDNA sequence of CBA mouse strain.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF367720; AAL09566.1; -.
DR MGD; MGI_96828; Lrp1.
GO; GO:0016020; C-membrane; IEA.
GO; GO:0005509; F:calcium ion binding; IEA.
GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000152; ASX_Hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00008; EGF_14.
DR Pfam; PF00057; ldl_recept_a; 31.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLA; 31.
DR SMART; SM00135; LY; 35.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 29.
KW EGF-like domain; Receptor.
SQ SEQUENCE 4545 AA; 504730 MW; 9904CF5DF5EB333E CRC64;

Query Match      96.0%; Score 486; DB 11; Length 4545;
Best Local Similarity 95.3%; Pred. No. 2.9e-51;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 26 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 85

QY 61 CVPMSRLCNGVQDCWMDGSDGPHCRE 86
DB 86 CVPMSRLCNGVQDCWMDGSDGPHCRE 111

RESULT 4
Q920Y4 PRELIMINARY; PRT; 4545 AA.
AC Q920Y4;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129P3/J;
RA Roebroek A.J.M.;

```


"Molecular cloning and sequencing of the Arawine L2M-receptor cDNA.";
 RL Biochim. Biophys. Acta 1173:71-74(1993).
 DR EMBL; X67469; CAA47817.1; -.
 DR PIR; S25111; S25111.
 DR HSP; Q07954; ICR8.
 DR MGD; MGI:96828; Lrp1.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR001881; EGF_CA.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000033; LDL_receptor_rep.
 DR Pfam; PF00008; EGF_3.
 DR Pfam; PF00057; ldl_recept_a; 31.
 DR Pfam; PF00058; ldl_recept_b; 34.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00179; EGF_CA; 3.
 DR SMART; SM00192; LDLA; 31.
 DR SMART; SM00135; LY; 33.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS00022; EGF_1; 5.
 DR PROSITE; PS01186; EGF_2; 8.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR PROSITE; PS01209; LDLRA_1; 26.
 DR PROSITE; PS00088; LDLRA_2; 29.
 DR EGF-like domain; Receptor.
 KW EGF-like domain; Receptor.
 SQ SEQUENCE 4545 AA; 504754 MW; A9263A81CF45E56D CRC64;
 Query Match 96.0%; Score 486; DB 11; Length 4545;
 Best Local Similarity 95.3%; Pred. No. 2.9e-51;
 Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KTCSFKQACRQDITCISKWRCDGERDCPDGSDAPEICPSKQACRCOPNEHNCGLGTEL 60
 DB 26 KTCSFKQACRQDITCISKWRCDGERDCPDGSDAPEICPSKQACRCOPNEHNCGLGTEL 85
 QY 61 CVPMRLCNGVQDCMGSDGEGPHCRE 86
 DB 86 CVPMRLCNGIQCDCMGSDGEGAHCRE 111
 RESULT 6
 ID Q8WY31 PRELIMINARY; PRT; 881 AA.
 AC Q8WY31;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Candidate tumor suppressor protein (fragment).
 GN LRP1B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20487554; PubMed=11031110;
 RA Liu C.-X., Musco S., Lisitsyna N.M., Yaklichkin S.Y., Lisitsyn N.A.;
 RT "Genomic organization of a new candidate tumor suppressor gene,
 LRP1B";
 RL Genomics 59:271-274(2000).
 DR EMBL; AF283342; AAL38107.1; -.
 DR EMBL; AF283327; AAL38107.1; JOINED.
 DR EMBL; AF283328; AAL38107.1; JOINED.
 DR EMBL; AF283329; AAL38107.1; JOINED.
 DR EMBL; AF283330; AAL38107.1; JOINED.
 DR EMBL; AF283331; AAL38107.1; JOINED.
 DR EMBL; AF283332; AAL38107.1; JOINED.
 DR EMBL; AF283333; AAL38107.1; JOINED.
 DR EMBL; AF283334; AAL38107.1; JOINED.
 DR EMBL; AF283335; AAL38107.1; JOINED.

DR EMBL; AF283336; AAL38107.1; JOINED.
 DR EMBL; AF283337; AAL38107.1; JOINED.
 DR EMBL; AF283338; AAL38107.1; JOINED.
 DR EMBL; AF283339; AAL38107.1; JOINED.
 DR EMBL; AF283340; AAL38107.1; JOINED.
 DR EMBL; AF283341; AAL38107.1; JOINED.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0005634; C-nucleus; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR001881; EGF_CA.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000033; LDL_receptor_rep.
 DR Pfam; PF00008; EGF_3.
 DR Pfam; PF00057; ldl_recept_a; 3.
 DR Pfam; PF00058; ldl_recept_b; 7.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00192; LDLA; 3.
 DR SMART; SM00135; LY; 9.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01209; LDLRA_1; 2.
 DR PROSITE; PS00068; LDLRA_2; 3.
 DR PROSITE; PS00037; MYB_1; 1.
 DR EGF-like domain.
 KW EGF-like domain.
 FT NON TER 881
 SQ SEQUENCE 881 AA; 99936 MW; 8325F9D984CADE5A CRC64;
 Query Match 55.3%; Score 280; DB 4; Length 881;
 Best Local Similarity 53.6%; Pred. No. 2.7e-26;
 Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;
 QY 3 CSPKQACRQDITCISKWRCDGERDCPDGSDAPEICPSKQACRCOPNEHNCGLGTEL 62
 DB 32 CDPGEFLCHDHTCVSQSWLDCGDPDSDSDSLDTCPEVEIKPLNHIACLGTKCV 91
 QY 63 PMSRLCNGVQDCMGSDGEGPHCRE 86
 DB 92 HSLQCNGLVDCPDGDEGVHCQE 115
 RESULT 7
 ID Q9NZR2 PRELIMINARY; PRT; 4599 AA.
 AC Q9NZR2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Low density lipoprotein receptor related protein-deleted in
 DE tumor.
 GN LRPDIT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20227253; PubMed=10766186;
 RA Liu C.-X., Musco S., Lisitsyna N.M., Forgacs E., Minna J.D.,
 RA Lisitsyn N.A.;
 RT "LRP-DIT, a putative endocytic receptor gene, is frequently
 RT inactivated in non-small cell lung cancer cell lines";
 RL Cancer Res. 60:1961-1967(2000).
 DR EMBL; AF176832; AAF70379.1; -.
 DR HSP; Q07954; ICR8.
 DR Genew; HGNC:6693; LRP1B.
 DR GO; GO:0015029; F:internalization receptor activity; TAS.
 DR GO; GO:0015031; P:protein transport; TAS.

GO; GO:0006988; P:receptor mediated endocytosis; TAS.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00008; EGF; 15.
DR Pfam; PF00057; ldl_recept_a; 32.
DR Pfam; PF00058; ldl_recept_b; 34.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00192; LDLA; 32.
DR PROSITE; PS00010; ASX HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 32.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00139; THIOI PROTEASE_CYS; 1.
KW EGF-like domain; Lipoprotein; Receptor.
SQ SEQUENCE 4599 AA; 515498 MW; 9A97A555F1FAAE66 CRC64;

Query Match 55.3%; Score 280; DB 4; Length 4599;
Best Local Similarity 53.6%; Pred. No. 1.5e-25;
Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAQRCPQNEHNCGLGTCLCV 62
Db 32 CDGFEFLCHDHVTCVQSQSWLDCGDPDPCDDSDLSLTCPEVEIKPLNHIACLGKNCV 91

QY 63 PMSRLNGVQDCMDGSDGPHCRE 86
Db 92 HLSQLCNGVLDCPDGDEGVHCQE 115

RESULT 8
Q9J118 PRELIMINARY; PRT; 4599 AA.
ID Q9J118
AC Q9J118
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Low density lipoprotein receptor related protein LRP1B/LRP-DIT.
GN LRP1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Yaklichkin S., Lisitsyn N.;
RT "The structure of the mouse homologue of the human candidate tumor
suppressor gene LRP1B/LRP-DIT.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AF270884; AAF81758.1; -
HSP; Q07954; ICB8
MD; MGI-2151136; Lrp1b.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR00152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00008; EGF; 15.

DR Pfam; PF00057; ldl_recept_a; 32.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00192; LDLA; 32.
DR PROSITE; PS00010; ASX HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 32.
DR PROSITE; PS00139; THIOI PROTEASE_CYS; 1.
KW EGF-like domain; Lipoprotein; Receptor.
SQ SEQUENCE 4599 AA; 513623 MW; 11462A3354FFB200 CRC64;

Query Match 53.4%; Score 270; DB 11; Length 4599;
Best Local Similarity 51.2%; Pred. No. 2.6e-24;
Matches 43; Conservative 13; Mismatches 28; Indels 0; Gaps 0;

QY 3 CSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAIPCQSKAQRCPQNEHNCGLGTCLCV 62
Db 32 CDGFEFLCHDHVTCVQSQSWLDCGDPDPCDDSDLSLTCPEVEIKPLNHIACLGSSACV 91

QY 63 PMSRLNGVQDCMDGSDGPHCRE 86
Db 92 HLSKLCNGVDCPDGDEGVHCQE 115

RESULT 9
O46131 PRELIMINARY; PRT; 883 AA.
ID O46131
AC O46131
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipophorin receptor precursor.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fat body;
RA Dantuma N.P., Potters M., Bogerd J., Vanderhorst D.J.;
RT "An insect homolog of the vertebrate very low density lipoprotein
receptor mediates endocytosis of lipophorins.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
EMBL; AJ000010; CAA03855.1; -
HSP; P01130; ILDR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 6.
DR PROSITE; PS00068; LDLRA_2; 8.
KW EGF-like domain; Receptor; Signal.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 883 LIPOPHORIN RECEPTOR.

SQ SEQUENCE 883 AA; 98332 MW; 86D0056AD80A05B CRC64;

Query Match 41.3%; Score 209; DB 5; Length 883;
Best Local Similarity 46.2%; Pred. No. 2e-17; Indels 2;
Matches 36; Conservative 12; Mismatches 26; Gaps 4;

QY 3 CSPKQAFACDQITCISKGWRCDCGERCPDGSDEAPEICPQSKAQRCQPNHNC-LGTEL 62
DB 255 CUPREFECLDRWTCIHQSWVCDGDRCPDGSDEVDSCRHN---MTCRPDQFC-RNRICI 310
QY 63 PMSRLCNGVQDCWDGSD 80
DB 311 PGLHCSGHADCSGSD 328

RESULT 10
Q9VBN1 PRELIMINARY; PRT; 883 AA.

AC Q9VBN1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CG31094-PA.
GN C34861.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blag J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jialali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., McBarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Ananatides P.G., Brandon R.C., Rogers Y.,
RA Bazon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jialali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA Ibegwan C., Jialali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochink S.E., Smith C.D.,
RA Tuzy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., S.E.;
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003753; AAF56498.2; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR008210; ISGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00057; ldl_recept_a; 7.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 7.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 7.
SQ SEQUENCE 883 AA; 98319 MW; 8245B1A3E9393F0 CRC64;

Query Match 41.0%; Score 207.5; DB 5; Length 883;
Best Local Similarity 50.0%; Pred. No. 3.1e-17;
Matches 40; Conservative 11; Mismatches 24; Indels 5; Gaps 3;

QY 2 TCSPKQAFACDQITCISKGWRCDCGERCPDGSDEAPEICPQSKAQRCQPNHNC-LGTEL 60
DB 30 TCSSDQFRCNG-NCIPNKRCDQESDCADGSDANELC---RARTCSPDEYACKSGEG 85
QY 61 CVPMSRLCNGVQDCWDGSD 80
DB 86 CVPLAWMCDQSKDCGSDGSD 105

RESULT 11
Q9V383 PRELIMINARY; PRT; 4699 AA.
ID Q9V383
AC Q9V383

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2006 (TrEMBLrel. 25, Last annotation update)
DE CG8706 protein.
GN CG33087 OR CG8706 OR CG33068.
OS Drosophila melanogaster (fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX STRAIN=Berkely.
RP SEQUENCE FROM N.A.
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arlil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter A.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherty J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flesler C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Snie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Fierriera S., Frise E., Galle R.P., Garq N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genomes.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA M.Sra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,

Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Flybase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DE EMBL; AEO03839; AAF59114.3; -
DR HSSP; P01130; 1AUJ.
DR FlyBase; FBgn0053087; CG33087.
DR GO; GO:0015020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti.; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR00152; Asx_hydroxyl_s.
DR InterPro; IPR001881; EGF_Ga.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_reg.
DR Pfam; PF00008; EGF_12.
DR Pfam; PF00057; ldl_recept_a; 27.
DR Pfam; PF00058; ldl_recept_b; 27.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 39.
DR SMART; SM00179; EGF_CA; 13.
DR SMART; SM00192; LDLA; 31.
DR SMART; SM00135; LY; 31.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 22.
DR PROSITE; PS00068; LDLRA_2; 29.
KW EGF-like domain.
SQ SEQUENCE 4699 AA; 524916 MW; 166D46B38C74A6BF CRC64;
Query Match 40.6%; Score 205.5; DB 5; Length 4699;
Best Local Similarity 45.8%; Pred. No. 3e-16;
Matches 38; Conservative 9; Mismatches 27; Indels 9; Gaps 3;
QY 2 TCSPKQAFACRDQITCISKWECDCGERDCPDGSDPEAICPOSKAQRCQPNHNCUGTCLC 61
Db 2756 SCDFGQFFRCASG-NCIAGSWHCDCGKDCPDGSDPEI-----NCRTECHNQFAC--DKTC 2806
QY 62 VPMRLCNGVQDCMDGSDGPHC 84
Db 2807 IPASWQCDCGKSDCEDGSDGPHC 2829
RESULT 12
QYW57
ID Q7YW57 PRELIMINARY; PRT; 891 AA.
AC Q7YW57;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipophorin receptor.
OS Aedes aegypti (yellow fever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
ON NCBI_TaxID=7159;

```

RN      SEQUENCE FROM N.A.
RP      Seq S.-J., Jun H.-M., Sun J., Sappington T.W., Raikhel A.S.;
RT      "tissue- and stage-specific expression of two lipophorin receptor
RT      variants with seven and eight ligand-binding repeats in the adult
RT      mosquito.";
RL      Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AY348869; AAL16410.1; -.
KW      Receptor.
SQ      SEQUENCE. 891 AA; 99303 MW; 82D411CD45EB271A CRC64;

Query Match      40.4%; Score 204.5; DB 5; Length 891;
Best Local Similarity 44.9%; Pred. No. 7.3e-17;
Matches 35; Conservative 11; Mismatches 29; Indels 3; Gaps 1;

QY      3 CSPKQACRDIQICISKGRCDGRCPCDGSDEAPICPSKAQRCPNHNCLGTCLCV 62
DB      205 CLSLEYQCSRITCIHKSWICDGEKDCPQGDDEMPIC---QNVTCRPDQFQCKDKTCI 261

QY      63 PMSRLCNGVQDCMDGSD 80
DB      262 NGHFHCNGKPECSDGSD 279

RESULT 13
Q963T3      PRELIMINARY; PRT; 1156 AA.
AC      Q963T3
DT      01-DEC-2001 (T-EMBLrel. 19, Created)
DT      01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      Lipophorin receptor.
OS      Aedes aegypti (Yellowfever mosquito).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX      NCBI_TaxID=7159;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21272543; PubMed=11378410;
RA      Cheon H., Seo S., Sun J., Sappington T.W., Raikhel A.S.;
RT      "Molecular characterization of the VLDR receptor homolog mediating
RT      binding of lipophorin in oocyte of the mosquito Aedes aegypti.";
RL      Insect Biochem. Mol. Biol. 31:753-760(2001).
DR      EMBL: AF355595; AAK72954.1; -.
DR      GO: 0016020; C:membrane; IEA.
DR      GO: 0005509; F:calcium ion binding; IEA.
DR      GO: 0004872; F:receptor activity; IEA.
DR      InterPro: IPR00152; A:EGF-like.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR006209; EGF-like.
DR      InterPro: IPR002172; LDL_receptor_A.
DR      InterPro: IPR000033; LDL_receptor_rep.
DR      Pfam: PF00057; ldl_recept_a; 8.
DR      PRINTS: PR00261; LDLRECEPTOR.
DR      SMART: SM00179; EGF_CA; 1.
DR      SMART: SM00192; LDLRA_2; 8.
DR      SMART: SM00135; LY; 5.
DR      PROSITE: PS00010; ASX_HYDROXYL; 1.
DR      PROSITE: PS01186; EGF_2; 1.
DR      PROSITE: PS01209; LDLRA_1; 7.
DR      PROSITE: PS50068; LDLRA_2; 8.
KW      Receptor.
SQ      SEQUENCE. 1156 AA; 128931 MW; 050D55E6C860E30F CRC64;

Query Match      40.4%; Score 204.5; DB 5; Length 1156;
Best Local Similarity 44.9%; Pred. No. 9.6e-17;
Matches 35; Conservative 11; Mismatches 29; Indels 3; Gaps 1;

QY      3 CSPKQACRDIQICISKGRCDGRCPCDGSDEAPICPSKAQRCPNHNCLGTCLCV 62
DB      299 CLSLEYQCSRITCIHKSWICDGEKDCPQGDDEMPIC---QNVTCRPDQFQCKDKTCI 355

QY      63 PMSRLCNGVQDCMDGSD 80
DB      262 NGHFHCNGKPECSDGSD 279

QY      63 PMSRLCNGVQDCMDGSD 80
DB      356 NGHFHCNGKPECSDGSD 373

RESULT 14
Q8WY29      PRELIMINARY; PRT; 1537 AA.
AC      Q8WY29
DT      01-MAR-2002 (T-EMBLrel. 20, Created)
DT      01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      Candidate tumor suppressor protein (fragment).
GN      LRP1B.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20487554; PubMed=11031110;
RA      Liu C.-X., Musco S., Lisitsyna N.M., Yaklichkin S.Y., Lisitsyn N.A.;
RT      "Genomic organization of a new candidate tumor suppressor gene,
RT      LRP1B";
RL      Genomics 69:271-274(2000).
DR      EMBL: AF283408; AAL38109.1; -.
DR      EMBL: AF283376; AAL38109.1; JOINED.
DR      EMBL: AF283377; AAL38109.1; JOINED.
DR      EMBL: AF283378; AAL38109.1; JOINED.
DR      EMBL: AF283379; AAL38109.1; JOINED.
DR      EMBL: AF283380; AAL38109.1; JOINED.
DR      EMBL: AF283381; AAL38109.1; JOINED.
DR      EMBL: AF283382; AAL38109.1; JOINED.
DR      EMBL: AF283383; AAL38109.1; JOINED.
DR      EMBL: AF283384; AAL38109.1; JOINED.
DR      EMBL: AF283385; AAL38109.1; JOINED.
DR      EMBL: AF283386; AAL38109.1; JOINED.
DR      EMBL: AF283387; AAL38109.1; JOINED.
DR      EMBL: AF283388; AAL38109.1; JOINED.
DR      EMBL: AF283389; AAL38109.1; JOINED.
DR      EMBL: AF283390; AAL38109.1; JOINED.
DR      EMBL: AF283391; AAL38109.1; JOINED.
DR      EMBL: AF283392; AAL38109.1; JOINED.
DR      EMBL: AF283393; AAL38109.1; JOINED.
DR      EMBL: AF283394; AAL38109.1; JOINED.
DR      EMBL: AF283395; AAL38109.1; JOINED.
DR      EMBL: AF283396; AAL38109.1; JOINED.
DR      EMBL: AF283397; AAL38109.1; JOINED.
DR      EMBL: AF283398; AAL38109.1; JOINED.
DR      EMBL: AF283399; AAL38109.1; JOINED.
DR      EMBL: AF283400; AAL38109.1; JOINED.
DR      EMBL: AF283401; AAL38109.1; JOINED.
DR      EMBL: AF283402; AAL38109.1; JOINED.
DR      EMBL: AF283403; AAL38109.1; JOINED.
DR      EMBL: AF283404; AAL38109.1; JOINED.
DR      EMBL: AF283405; AAL38109.1; JOINED.
DR      EMBL: AF283406; AAL38109.1; JOINED.
DR      EMBL: AF283407; AAL38109.1; JOINED.
DR      GO: 0016020; C:membrane; IEA.
DR      GO: 0005509; F:calcium ion binding; IEA.
DR      InterPro: IPR000152; ASX_HYDROXYL_S.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR006209; EGF-like.
DR      InterPro: IPR000033; LDL_receptor_A.
DR      InterPro: IPR002172; LDL_receptor_rep.
DR      Pfam: PF00057; ldl_recept_a; 18.
DR      Pfam: PF00058; ldl_recept_b; 9.
DR      PRINTS: PR00261; LDLRECEPTOR.
DR      SMART: SM00179; EGF_CA; 3.
DR      SMART: SM00192; LDLRA_1; 18.
DR      SMART: SM00135; LY; 8.
DR      PROSITE: PS00010; ASX_HYDROXYL; 3.
DR      PROSITE: PS01186; EGF_2; 2.
```

DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 14.
DR PROSITE; PS50068; LDLRA_2; 18.
KW EGF-like domain.
FT NON_TER 1
FT NON_TER 1537 1537
SQ SEQUENCE 1537 AA; 172525 MW; 6F38AAFC818F135C CRC64;

Query Match
Best Local Similarity 40.1%; Score 203; DB 4; Length 1537;
Matches 41; Conservative 7; Mismatches 27; Indels 4; Gaps 2;
QY 2 TCSPKQFACRQDITCISKWRCGDERDCPDGSDPEAICPQSKAQCQPNEHNCILGTEL 61
DB 43 SCWNQFACSAQ-KCISKWICGDEDDCGDLDESISIC---CAITCAADMFSCQGSRA 98
QY 62 VPMRLCNGVQDCMDGSDSDE 80
DB 99 VPRHWLDCGERDCPDGSDSDE 117

RESULT 15

Q81GR9 PRELIMINARY; PRT; 739 AA.
AC Q81GR9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R33884P.
GN CG4861.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokston P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Munsall C.J., Nuroo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT001632; AAN71387.1; -.
DR FlyBase; FBgn039363; CG4861.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00009; EGF; 1.
DR Pfam; PF00057; ldl_recept_a; 5.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00132; LDL; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 5.
DR PROSITE; PS00068; LDLRA_2; 5.
SQ SEQUENCE 739 AA; 82203 MW; 53162F0A9EF82823 CRC64;

Query Match
Best Local Similarity 39.9%; Score 202; DB 5; Length 739;
Matches 35; Conservative 11; Mismatches 28; Indels 4; Gaps 2;

QY 3 CSPKQFACRQDITCISKWRCGDERDCPDGSDPEAICPQSKAQCQPNEHNCILGTEL 62
DB 105 CLSHEYQCKDRITCLHHSWLCGDRDCPDGDEHTANC---KNVTCTADQFQC-GDRSCI 160
QY 63 PMSRLCNGVQDCMDGSDSDE 80
DB 161 PGHLTCNGDKDCADGSDSDE 178
Search completed: September 17, 2004, 11:17:54
Job time : 119 secs

This Page Blank (uspto)